SANDIA REPORT
SAND97–2133 • UC–405
Unlimited Release
Printed September 1997

Scaling Vectors and Multiwavelets in Numerical Differential Equations–Some Approximation-Theoretic and Numerical Issues

Peter R. Massopust

Prepared by
Sandia National Laboratories
Albuquerque, New Mexico 87185 and Livermore, California 94550

Sandia is a multiprogram laboratory operated by Sandia Corporation, a Lockheed Martin Company, for the United States Department of Energy under Contract DE-AC04-94AL85000.

Approved for public release; further dissemination unlimited.

Sandia National Laboratories

Issued by Sandia National Laboratories, operated for the United States Department of Energy by Sandia Corporation.

Printed in the United States of America. This report has been reproduced directly from the best available copy.

Available to DOE and DOE contractors from
Office of Scientific and Technical Information
PO Box 62
Oak Ridge, TN 37831

Prices available from (615) 576-8401, FTS 626-8401

Available to the public from
National Technical Information Service
US Department of Commerce
5285 Port Royal Rd
Springfield, VA 22161

NTIS price codes
Printed copy: A08
Microfiche copy: A01

SAND97-2133
Unlimited Release
Printed September 1997

Distribution
Category UC-405

Scaling Vectors and Multiwavelets in Numerical Differential Equations – Some Approximation-theoretic and Numerical Issues[1]

Peter R. Massopust
Applied and Numerical Mathematics Department
Sandia National Laboratories
P.O. Box 5800
Albuquerque, NM 87185-1110
U.S.A.

Abstract

This report summarizes some of the approximation-theoretic and numerical issues encountered in solving operator equations of the form $Lu = f$. Particular emphasis is placed on Galerkin and finite element approximations using multiwavelets. Examples are used to illustrate some of the issues.

[1]Keywords and Phrases: Scaling vector, refinability, Sobolev spaces, elliptic operators, preconditioning, Jackson- and Bernstein-type inequalities, multiwavelets, multiresolution analysis.

AMS Classification (1991 Revision): 41A30, 41A17, 35J25,46E35, 65N30.

1 Introduction

Galerkin approximations and finite element methods for operator equations of the form $Lu = f$ play an important role in the theory of numerical differential equations. Both are based on ascending sequences of trial spaces generated, in general, by hierarchical bases. The different discretization levels are allowed to interact, and this interaction is then utilized to reduce errors in the numerical approximations. The use of local bases gives rise to sparse stiffness matrices but these matrices are ill-conditioned due to the interaction between the different discretization levels. The condition numbers increase exponentially in the number of grid points or, equivalently, with higher approximation levels. Preconditioners have to be used to exploit the full potential of iterative solution methods. The theory of hierarchical preconditioners for Galerkin and finite element methods is well developed, see for instance [2, 7, 15, 17], and such preconditioners are explicitly known. The construction of the preconditioner relies on the properties of the operator L and certain estimates established in a scale of Sobolev spaces to be used in the Galerkin approximations of the solution u.

Scaling vectors are collections of special functions that may be used to construct nested sequences of approximation or trial spaces, called a *multiresolution analysis.* Associated with a scaling vector is a *multiwavelet* which generates the *difference spaces*, usually referred to as *multiwavelet spaces*, between successive approximation spaces. Certain classes of scaling vectors and their associated multiwaveles provide fast *multiscale transformations* between the different levels of discretization. Loosely speaking, in such approximation schemes the scaling vector is the carrier of the coarse level information whereas the multiwavelet carries the detail or fine structure information. Scaling vectors and multiwavelets have the potential to provide numerically and computationally efficient algorithms for Galerkin-type approximations of operator equations. In particular, they seem to give the correct framework for adaptive and multiresolution schemes to obtain solutions that vary drastically in space and time and develop singularities. Moreover, they fit into well-established multiscale methods for operator equations and possess sound approximation-theoretic foundations.

The structure of the survey report is as follows. In Section 2 the mathematical setting for the type of operator equation considered in this report is introduced and some remarks are made about the existence of unique solutions. Next, scaling vectors and multiwavelets are introduced as generators of *refinable spaces*. Some properties are discussed and fast multiscale transformations are introduced. A specific scaling vector and multiwavelet, the so-called *GHM* and *DGHM elements*, are presented. As the GHM and the DGHM element is based on *fractal functions*, a few brief remarks are made about the properies and features of such functions. The *recursive structure* of the GHM and DGHM element, its *interpolatory nature*, and its property of being easily adjustable to both *bounded intervals* (without introducing new boundary functions) and *nonuniform geometries* are direct consequences of its construction. The next section focuses on the approximation-theoretic foundations of multiscale methods. In particular, for differential operator equations of the form $Lu = f$ the relevant Sobolev spaces are defined and the fundamental Jackson- and Bernstein-type estimates for multiscale methods are presented. The convergence and approximation properties of scaling vectors and multiwavelets then follow from these estimates. Furthermore, the preconditioner for the Galerkin method is derived, and it is shown that it may be interpreted as a change of basis operator between different levels of approximations. Section 4 considers a Galerkin method for a very simple differential operator equation on $[0, 1]$, namely, $L = -\Delta$. Here, for illustrative purposes, all relevant quantities are explicitly derived.

2 Refinable Spaces, Scaling Vectors, and Multiwavelets

In this section the mathematical setting for operator equations of the form $Lu = f$ is presented, and the important concept of a *refinable space* is introduced. The generator of this space, the so-called scaling vector, is presented next. The associated multiwavelet is then defined as the generator of certain refinable difference spaces. Properties of these functions are discussed and a particular example of a scaling vector and its associated multiwavelet considered. Some brief remarks about fractal functions and their features close out the section.

2.1 The mathematical setting

This report focuses on some of the approximation-theoretic and numerical issues regarding the solutions of operator equations of the form $Lu = f$. For the sake of simplicity, it may be assumed that L is a second order elliptic differential operator, although the results stated here will apply to a much larger class of operators. To obtain an approximate solution to such an operator equation several points need to be addressed.

- A function space $\mathcal{F}$ must be identified in which the solution u is expected to lie. In most cases $\mathcal{F}$ may be assumed to be a *Hilbert space.*
- Conditions need to be imposed on the operator L and possibly the right-handside f that guarantee the existence of a *unique solution* in $\mathcal{F}$.
- A basis of $\mathcal{F}$ is to be found so that an *algebraic system* associated with $Lu = f$ is efficiently solvable.

These issues will now be presented in a more precise fashion.

Operator equations of the form $Lu = f$ have a natural setting in the theory of Hilbert spaces. Let $\mathcal{H}$, $\mathcal{H}_1$, and $\mathcal{H}_2$ by Hilber spaces satisfying either one of the following two sets of inclusions

$$\mathcal{H}_1 \subseteq \mathcal{H} \subseteq \mathcal{H}_2 \quad \text{or} \quad \mathcal{H}_2 \subseteq \mathcal{H} \subseteq \mathcal{H}_1. \tag{2.1}$$

The above inclusions, called *continuous embeddings*, have to be interpreted in the following way. Denote by $\langle \cdot, \cdot \rangle_{\mathcal{F}}$, where $\mathcal{F}$ is either $\mathcal{H}$, or $\mathcal{H}_1$, or $\mathcal{H}_2$, the *inner product* on the Hilbert space $\mathcal{F}$, and by $\| \cdot \|_{\mathcal{F}}^2 = \langle \cdot, \cdot \rangle_{\mathcal{F}}$ the associated *norm.* Then Eqn. (2.1) means that there exist positive constants c_1 and c_2 so that

$$\boxed{c_1 \| \cdot \|_{\mathcal{H}_2} \leq \| \cdot \|_{\mathcal{H}} \leq c_2 \| \cdot \|_{\mathcal{H}_1},} \tag{2.2}$$

and similarly for the second set of inclusions. For most operator equations of the form considered here, $\mathcal{H}_2$ is the *dual space* of $\mathcal{H}_1$, i.e., the linear space of all *continuous linear functionals* $\varphi : \mathcal{H}_1 \to \mathbb{R}$.

Example 2.1 *The differential equation*

$$-u'' = f, \quad \text{on } [0,1] \text{ with } u(0) = u(1) = 0 \tag{2.3}$$

has the weak *formulation*

$$-\int_0^1 u''vdx = \int_0^1 fvdx$$

Integration by parts yields:

$$\int_0^1 u'\,v'dx - (u'(1)v(1) - u'(0)v(0)) = \int_0^1 fvdx. \tag{2.4}$$

The last equation suggests that the solution u should be chosen from the Sobolev space $H^1([0,1])$, *i.e., the space of all functions that are limits of* infinitely differentiable functions ϕ with compact support [2] *in* $(0,1)$ *relative to the norm* $\|\phi\|_{H^1} = \|\phi\|_{L^2} + \|\phi'\|_{L^2}$, *and v from the Sobolev space* $H_0^1([0,1]) = \{\phi \in H^1([0,1]) : \phi(0) = \phi(1) = 0\}$. *(Note that since u needs to satisfies the boundary conditions $u(0) = u(1) = 0$, u is also in $H_0^1([0,1])$.) Moreover, since L here is the differential operator d^2/dx^2, f can be chosen from the space $H^{-1}([0,1])$, the dual space of H_0^1. (The Sobolev space $H^{-1}([0,1])$ contains* distributions *such as the Dirac δ-function.) Since it is known that*

$$\boxed{H_0^1([0,1]) \subseteq L^2([0,1]) \subseteq H^{-1}([0,1]),} \tag{2.5}$$

one finally has $\mathcal{H} = L^2([0,1])$. The weak formulation (2.4) may then be expressed in the form

$$\langle u', v'\rangle_{L^2} = \langle f, v\rangle_{L^2}. \tag{2.6}$$

Remark 2.1 *In the above example, the differential operator $L = -d^2/dx^2$ has order two and therefore, if $u \in H^1([0,1])$ then $f = Lu \in H^{1-2}([0,1]) = H^{-1}([0,1])$. More generally, if L is an operator of order t and if u is a function in the Sobolev space H^s then Lu is in the Sobolev space H^{s-2t}. Here, for any positive integer s, $H^s(\Omega)$ is the space of all functions that are limits of infinitely differentiable functions ϕ with compact support in Ω relative to the norm $\|\phi\|_{H^1} = \|\phi\|_{L^2} + \sum_{n=1}^s \|f^{(s)}\|_{L^2}^2$. For latter purposes, Sobolev spaces with noninteger index s need to be defined. One way of doing this is via* Fourier transforms. *To this end, let $s > 0$ and nonintegral,*

$$\boxed{H^s(\mathbb{R}) = \left\{f \in L^2(\mathbb{R}) : \int_{\mathbb{R}} (1+|\omega|^2)^s \, |\hat{f}(\omega)|^2 \, d\omega < \infty\right\}.} \tag{2.7}$$

[2]The support of a function ϕ is the largest closed interval outside of which the function is identically equal to zero.

Definition (2.7)) may be then be used to define the spaces $H^s(\Omega)$. Sobolev spaces $H^s(\Omega)$ with negative index s are the dual spaces of $H^{-s}(\Omega)$.

The question of *existence of a unique solution of $Lu = f$* is addressed next. If the operator is a bounded linear one-to-one and onto mapping from the Hilbert space $\mathcal{H}_1$ into the Hilbert space $\mathcal{H}_2$, then the aforementioned operator equation has a unique solution. The conditions on L may be re-expressed in the following way: there exist two positive constants c_1 and c_2 such that

$$\boxed{0 < c_1 \|Lu\|_{\mathcal{H}_2} \leq \|u\|_{\mathcal{H}_1} \leq c_2 \|Lu\|_{\mathcal{H}_2}, \quad \text{for all } u \in \mathcal{H}_1.} \tag{2.8}$$

These above conditions are for instance satisfied if L is a *linear self-adjoint positive definite*[3] *elliptic differential operator.* (Cf. also Example (2.1).)

Now the last point, the choice of an appropriate basis for the Hilbert space $\mathcal{F}$ containing the solution u, is taken up. To this end, let $\{\psi_\lambda : \lambda \in \Lambda\}$ be a countable basis for $\mathcal{F}$. (Here Λ is an index set which may be assumed to be a subset of the positive integers or the positive integers themselves.) Then each $u \in \mathcal{F}$ can be expressed in the form

$$u = \sum_{\lambda \in \Lambda} c_\lambda \psi_\lambda, \quad \text{for some real coefficients } c_\lambda. \tag{2.9}$$

The sum in this representation of u is interpreted as follows. Let Λ_k be a finite subset of Λ with the property that $\#\Lambda_k \to \#\Lambda$[4] as $k \to \infty$. Then Eqn. (2.9) means

$$\left\| u - \sum_{\lambda \in \Lambda_k} c_\lambda \psi_\lambda \right\|_{\mathcal{F}} \to 0 \quad \text{as } k \to \infty.$$

Substitution of (2.9) into the operator equation $Lu = f$ yields

$$\sum_{\lambda \in \Lambda} c_\lambda \, L\psi_\lambda = f. \tag{2.10}$$

[3]A linear operator L is called *self-adjoint* if $\langle Lu, v\rangle = \langle u, Lv\rangle$ and *positive definite* if $\langle Lu, u\rangle > 0$.

[4]The number of elements in an index set Λ is denoted by $\#\Lambda$.

Letting $\boldsymbol{c}$ and $\boldsymbol{\psi}$ be the (column) vector $(c_\lambda : \lambda \in \Lambda)$ and $(\psi_\lambda : \lambda \in \Lambda)$, respectively, multiplying on the right by $\boldsymbol{\psi}$, and transposing (^{T} denotes the transpose) the resulting equation produces an *algebraic system* for the unknowns $\boldsymbol{c}$:

$$\boxed{\langle L\boldsymbol{\psi}, \boldsymbol{\psi}\rangle^T \boldsymbol{c} = \langle f, \boldsymbol{\psi}\rangle^T.} \tag{2.11}$$

At this point several options are available to obtain a numerical solution.

- Choose the basis $\{\psi_\lambda : \lambda \in \Lambda\}$ in such a way that the – possibly infinite – matrix $\langle L\boldsymbol{\psi}, \boldsymbol{\psi}\rangle^T$ is diagonal. This, however, is not possible in the majority of cases.

- Apply a *preconditioner* $\boldsymbol{P}$ to the system (2.11) such that

  $$\boldsymbol{P}\langle L\boldsymbol{\psi}, \boldsymbol{\psi}\rangle^T \boldsymbol{P}$$

  is efficiently solvable.

- Choose a basis in which $\langle L\boldsymbol{\psi}, \boldsymbol{\psi}\rangle^T$ has a *sparse* representation.

- Instead of using the entire basis $\{\psi_\lambda : \lambda \in \Lambda\}$, adaptively consider only those basis elements which give an accurate description of the solution u and discard the remaining ones. This amounts to solving the system in certain *subspaces* of $\mathcal{F}$.

The next subsection contains a brief summary of terminology and symbols used.

2.2 Frequently used symbols

$\mathbb{N}$: set of natural numbers $1, 2, \ldots$
$\mathbb{Z}$: set of integers $\ldots, -2, -1, 0, 1, 2, \ldots$
$\mathbb{R}$: set of real numbers or the real line

Ω : open connected domain on the real line $\mathbb{R}$ ($\Omega = \mathbb{R}$ included)

$\tau[\Phi]$: the linear span of all integer-translates of a function Φ: lin span$\{\Phi(x-\ell) : \ell \in \mathbb{Z}\}$. Any $f \in \tau[\Phi]$ is of the form $f(x) = \sum_{\ell=-\infty}^{+\infty} c_\ell \Phi(x-\ell)$.

$L^2(\Omega)$: space of all functions f with the property that $\int_\Omega |f(x)|^2 dx$ is finite (*square integrable functions on* Ω.)

$\|\cdot\|_{L^2}$: L^2 norm: $\|f\|_{L^2} = \sqrt{\int_\Omega |f(x)|^2 dx}$.

$\langle \cdot, \cdot \rangle$: L^2 inner product: $\langle f, g \rangle = \int_\Omega f(x) g(x) dx$

$\sigma[\Phi]$: the closure of $\tau[\Phi]$ in $L^2(\mathbb{R})$, i.e., the set of all functions f such that, given any sequence $\{\Phi_n\}$ of functions from $\tau[\Phi]$, the norm $\|f - \Phi_n\|_{L^2}$ can be made arbitrarily small by choosing n large enough.

$\widehat{f}$: Fourier transform of a function $f \in L^2(\mathbb{R})$:

$$\widehat{f}(\omega) = \int_{\mathbb{R}} e^{ix\omega} f(x)\, dx$$

$H^n(\Omega)$: Sobolev space consisting of all functions f defined on Ω with the property that $f^{(n)} = d^n f/dx^n \in L^2(\Omega)$ (n positive integer)

$\|\cdot\|_{H^n}$: Sobolev norm: $\|f\|_{H^n} = \|f\|_{L^2} + \sum_{\nu=1}^{n} \|f^{(n)}\|_{L^2}^2$ (n positive integer)

$H^s(\mathbb{R})$: Sobolev space of all functions $f \in L^2(\mathbb{R})$ such that $\int_{\mathbb{R}} (1+|\omega|^2)^s\, |\widehat{f}(\omega)|^2\, d\omega < \infty$. ($s > 0$)

$\|\cdot\|_{H^s}$: Sobolev norm: $\|f\|_{H^s} = \int_{\mathbb{R}} (1+|\omega|^2)^s\, |\widehat{f}(\omega)|^2\, d\omega$ ($s > 0$)

$H_0^s(\Omega)$: Sobolev space consisting of all functions $f \in H^s(\Omega)$ such that $f = 0$ on the boundary $\partial\Omega$ of Ω.

$\|\cdot\|_{\ell^2}$: norm for *square-summable sequences* $\{c_n\}$: $\|c_n\|_{\ell^2}^2 = \sum_{n=0}^{\infty} |c_n|^2$.

2.3 Shift-invariant spaces

A space of functions V contained in $L^2(\mathbb{R})$ is called *shift-invarinat* or *translation invariant* if for every function $f \in V$ and every $\ell \in \mathbb{Z}$ the *integer-translate* or *shift* $f(x-\ell)$ is also in V. Such shift-invariant spaces may be generated by choosing a particular function Φ, called *a generator*, and taking the linear span of all its integer-translates $\tau[h] = \text{lin span}\{h(x-\ell) : \ell \in \mathbb{Z}\}$. In order to ensure that the limits of sequences of functions from $\tau[\Phi]$ belong to $\tau[\Phi]$, the shift-invariant space needs to be the *closure* of $\tau[\Phi]$ in the L^2-norm.[5] The closure of $\tau[\Phi]$ is denoted by $\sigma[\Phi]$. A shift-invariant space V that is generated by a *single* function Φ is called a *principal shift-invariant space* and written as $V = \sigma[\Phi]$.

As an example of this procedure, consider the *hat function* defined by

$$h(x) = \begin{cases} 1-|x|, & -1 \le x \le 1 \\ 0 & \text{otherwise,} \end{cases} \tag{2.12}$$

[5]The closure of a space S in the L^2-norm consists of all functions f that are limits of sequences of functions from S.

Now take the linear span of all its integer-translates: $\tau[h] = \text{lin span}\{h(x - \ell) : \ell \in \mathbb{Z}\}$. Since the hat function h is in $L^2(\mathbb{R})$, $\int_{-1}^{1} |h(x)|^2 dx = 2/3$, the space $\sigma[h]$ consists of all functions f in $L^2(\mathbb{R})$ which can be written in the form

$$f(x) = \sum_{\ell=-\infty}^{+\infty} c_\ell h(x - \ell), \tag{2.13}$$

for a sequence of real coefficients satisfying

$$\|\{c_\ell\}\|_{\ell^2} = \sum_{\ell=-\infty}^{+\infty} |c_\ell|^2 < \infty. \tag{2.14}$$

It is worthwhile noting that sums such as those in Eqn. (2.13) which involve functions from $L^2(\mathbb{R})$ are to be understood in the L^2-sense, namely,

$$\boxed{\lim_{M,N\to\infty} \left\| f(x) - \textstyle\sum_{\ell=-M}^{+N} c_\ell h(x - \ell) \right\|_{L^2} = 0.} \tag{2.15}$$

The space $\sigma[h]$ may also be characterized as consisting of all *piecewise linear functions in $L^2(\mathbb{R})$ with integer knots.* In other words, any function f in $L^2(\mathbb{R})$ whose values $f(x)$ are given on integer knots $x = \ell \in \mathbb{Z}$ has a representation of the form (2.13).

Shift-invariant spaces may be constructed using more than one function. More precisely, if $\phi^1, \phi^2, \dots \phi^r$ are functions from $L^2(\mathbb{R})$, then the closure of the linear span of the translates of all these functions defines what is called a *finitely generated shift-invariant space, FIS space* for short. To simplify notation, let $\Phi = (\phi^1, \phi^2, \dots \phi^r)$ and write $V = \sigma[\Phi]$ instead of $V = \sigma[\{\phi^1, \phi^2, \dots \phi^r\}]$.

An example of a FIS space is given by $V = [h, q]$ where h is the hat function and $q(x) = \begin{cases} 4x(1-x), & 0 \le x \le 1 \\ 0, & \text{otherwise.} \end{cases}$ Since q is a quadratic function, the space $V = \sigma[h, q]$ contains all piecewise quadratic functions on intervals with integer endpoints.

2.4 Refinable spaces

It is an easy exercise to show that the hat function h satisfies the following equation:

$$h(x) = \frac{1}{2}h(2x+1) + h(2x) + \frac{1}{2}h(2x-1). \tag{2.16}$$

Equations of the form (2.16) are called *refinement or two-scale dilation equations*, for, they express a function, here h, in terms of the *dilates by two* and *integer translates* of itself. More generally, any equation of the form

$$\boxed{f(x) = \sum_{\ell \in \mathbb{Z}} g_\ell f(2x - \ell),} \tag{2.17}$$

is called a refinement equation. The sum in Eqn. (2.17) may be finite or infinite; it is finite if and only if the function f has compact support. Note that Eqn. (2.17) is equivalent to

$$\boxed{f(x/2) = \sum_{\ell \in \mathbb{Z}} g_\ell f(x - \ell).} \tag{2.18}$$

The concept of refinability gives rise to a *refinable space.* To this end, let D denote the operator which dilates by a factor of 2: $(Df)(x) = f(x/2)$. A space of functions V is called *refinable* provided that

$$D(V) \subset V \Longleftrightarrow f(x) \in V \quad \text{implies} \quad f(x/2) \in V. \tag{2.19}$$

An example of a refinable space is $V = \sigma[h]$. For, by Eqn. (2.16) with x replaced by $x/2$, $h(x/2)$ is in V. Thus, since every function f in $V = \sigma[h]$ is a linear combination of the translates of the hat function h, $f(x/2)$ is also in V. One can also interpret this result from an interpolation-theoretic point of view: if a function in $V = \sigma[h]$ interpolates at the integers. it also interpolates at the even integers.

A space $V = \sigma[\Phi]$, with generators $\Phi = \{\phi^1, \phi^2, \ldots \phi^r\}$ is refinable if and only if, the (column) vector of generators $\Phi = (\phi^1 \; \phi^2 \; \ldots \; \phi^r)^T$ satisfies

$$\boxed{\Phi(x) = \sum_{\ell \in \Lambda} \boldsymbol{G}_\ell \Phi(2x - \ell),} \tag{2.20}$$

for some sequence of real $r \times r$ matrices $\{\boldsymbol{G}_\ell : \lambda \in \Lambda\}$. Here Λ denotes a finite or possibly infinite subset of the integers $\mathbb{Z}$. As all generators $\phi^1, \ldots, \phi^r$ are elements of $L^2(\mathbb{R})$, i.e., $\Phi \in (L^2(\mathbb{R}))^r$, the matrices $\boldsymbol{G}_\ell$, $\ell \in \Lambda$, satisfy

$$\boxed{\sum_{\ell \in \Lambda} ||\boldsymbol{G}_\ell||^2 < \infty.} \tag{2.21}$$

Here $|| \cdot ||$ denotes a matrix norm, for instance, the spectral norm.

Consider again the principal shift-invariant space $V = \sigma[h]$ generated by the hat function h. Note that

$$\langle h, h(\cdot - 1\rangle = \int_{-1}^{1} h(x)h(x-1)dx = 1/6 = \langle h, h(\cdot + 1\rangle. \tag{2.22}$$

In other words, the generator h is *not* orthogonal[6] to its integer translates. As the hat function is used in finite element methods, this lack of orthogonality is reflected in the well-known fact that the stiffness matrix is tridiagonal. A way to remedy this situation is to replace the single generator h by a *pair* of generators that are orthogonal to each other as well as their integer translates. The following example shows how to obtain *orthogonal generators* for $V = \sigma[h]$.

Example 2.2 *Start again with the hat function $h(x)$. Introduce a new and yet unknown continuous function u supported on $[0,1]$, and define $V = \sigma[h, u]$.*

The main idea, due to [5], is to modify the hat function h in such a way that a new function v is obtained that satisfies

- *v is supported on $[-1,1]$;*
- *v is a linear combination of h, u, and $u(x+1)$;*
- *v is orthogonal to its translates $v(x \pm 1)$ and to u;*
- *$V = \sigma[u, v]$. (This means that the original space remains* unchanged*; only new generators have been chosen.)*

For this purpose, define

$$v(x) = h - \frac{\langle h, u\rangle}{\langle u, u\rangle} u - \frac{\langle h, u(x+1)\rangle}{\langle u, u\rangle} u(x+1). \tag{2.23}$$

Note that u and its translate $u(x+1)$ is projected out of the hat function, thsu making v orthogonal to u. As u is already orthogonal to its integer shifts, only v needs to be orthogonal to its translates (by symmetry it suffices to only consider the translate to the right)

$$\langle v, v \cdot -1\rangle = 0.$$

[6]Two functions in $L^2(\Omega)$ are called *orthogonal* if $\langle f, g\rangle_\Omega = \int_\Omega f(x)g(x)dx = 0$.

This, however, is equivalent to

$$\boxed{\langle h, h(\cdot - 1)\rangle = \tfrac{\langle v,u\rangle \langle v(\cdot-1),u\rangle}{\langle u,u\rangle}.} \tag{2.24}$$

Choosing for u the function

$$u(x) = \begin{cases} x, & 0 \le x \le 1, \\ 0, & \textit{otherwise} \end{cases}$$

yields

$$v(x) = \begin{cases} 1 - 3x/2, & 0 \le x \le 1, \\ 0, & \textit{otherwise} \end{cases}.$$

Clearly, u and v are orthogonal generators and since they are both supported on the unit interval, they are also orthogonal to their integer translates. Notice that the new generators u and v satisfy a matrix refinement equation*:*

$$\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} 1/2 & 0 \\ 3/4 & 1 \end{pmatrix} \begin{pmatrix} u(2x) \\ v(2x) \end{pmatrix}. \tag{2.25}$$

The generators u and v have one apparent disadvantage: they are not continuous *across integer intervals $[\ell, \ell + 1]$, $\ell \in \mathbb{Z}$.*

In order to construct generators for the space $V = \sigma[h]$ that are continuous across integer intervals, a more sophisticated choice for u must be made. To this end, employ the refinability condition. Namely, if $u(x/2) \in V$ then $u(x/2)$ must be a linear combination of $h(x-1)$, u, and $u(x-1)$:

$$\boxed{u(x/2) = h(x-1) + s_0 u(x) + s_1 u(x-1),} \tag{2.26}$$

for some real constants s_0 and s_1. But Eqn. (2.26) is recognized as a so-called *inhomogenous two-scale dilation equation.* The unique solutions of such equations are *affine fractal (interpolation) functions.* Since some of the properties of such functions are important for future developments, a short introduction to the theory of fractal functions is given in the next subsection.

Choosing $s_0 = s_1 =: s$ causes u to be symmetric about the line $x = 1/2$. Employing properties of affine fractal functions, one obtains

$$\begin{aligned} \langle h, u\rangle &= \langle h, u(x+1)\rangle = \frac{1}{4(1-s)} \\ \langle u, u\rangle &= \frac{2+s}{6(1-s)^2(1+s)} \end{aligned}$$

Since $\langle h, h(x-1)\rangle = \frac{1}{6}$, the orthogonality condition (2.24) yields

$$\boxed{s = -1/5}$$

The graphs of the two orthogonal generators are shown in Figure 1. In the

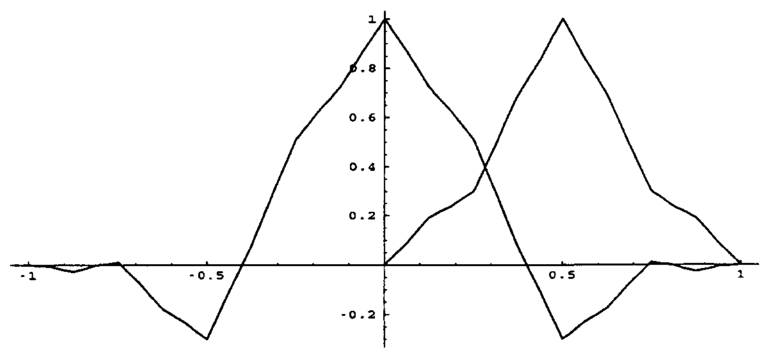

Figure 1: The orthogonal generators u and v.

literature, the functions u and v are referred to as the GHM element [9]. The function v consists of *two* affine fractal functions: one supported on $[-1,0]$ the other on $[0,1]$.

The FIS space $V^* = \sigma[\{u,v\}]$ generated by the GHM element contains the space $V = \sigma[h]$ generated by the hat function; for h is a linear combination of u, $u(x+1)$, and v (cf. Eqn. (2.23)). However, u is not in $V = \sigma[h]$; hence $V = \sigma[h]$ is a *proper* subspace of $V^* = \sigma[\{u,v\}]$. Nevertheless, the GHM element has the *same* approximation-theoretic properties as the hat function: piecewise linear and constant functions are reproduced exactly [6]. In other words, there exists a *finite* sequence of real vectors $\{\boldsymbol{c}_\ell\}$ such that

$$\boxed{x^p = \sum_\ell \boldsymbol{c}_\ell^T \begin{pmatrix} u(x-\ell) \\ v(x-\ell) \end{pmatrix}, \qquad p = 0, 1,} \tag{2.27}$$

on *any finite interval with integer endpoints.* The space of piecewise linear polynomials on integer intervals can therefore be described either as $V = \sigma[h]$ or as $V^* = \sigma[u, v]$.

2.5 Rudimentaries from the theory of fractal functions

In this subsection, an exemplatory introduction to the theory of fractal functions is presented. This presentation is not the most general one, but for the purposes of this report sufficient. The interested reader is referred to [12] for a more detailed and general introduction to the subject.

Let $\{(x_j, y_j) : j = 0, 1, 2\}$ be a given set of interpolation points and let f_0 be *any continuous function* satisfying

$$f_0(x_j) = y_j, \qquad j = 0, 1, 2.$$

Define an operator T by

$$\boxed{(Tf_0)(x) = \begin{cases} \lambda_0\left(\frac{x-b_0}{a_0}\right) + s_0 f_0\left(\frac{x-b_0}{a_0}\right), & x_0 \le x \le x_1 \\ \\ \lambda_1\left(\frac{x-b_1}{a_1}\right) + s_1 f_0\left(\frac{x-b_1}{a_1}\right), & x_1 < x \le x_2 \end{cases},} \tag{2.28}$$

where $\lambda_\ell(x) = c_\ell(x) + d_\ell$, $\ell = 0, 1$, is the *unique affine function* such that

$$(Tf_0)(x_0) = f_0(x_0), \qquad (Tf_0)(x_2) = f_0(x_2),$$
$$(Tf_0)(x_1-) = f_0(x_1) = (Tf_0)(x_1+).$$

Join-up conditions to guarantee continuity

The coefficients a_ℓ, b_ℓ, c_ℓ and d_ℓ are explicitly given by

$$\begin{matrix} a_\ell = \frac{x_{\ell+1}-x_\ell}{x_2-x_0}, & b_\ell = \frac{x_2 x_\ell - x_0 x_{\ell+1}}{x_2-x_0}, & \\ & & \ell = 0, 1 \\ c_\ell = \frac{y_{\ell+1}-y_\ell-s_\ell(y_2-y_0)}{x_2-x_0}, & d_\ell = \frac{x_2 y_\ell - x_0 y_{\ell+1} - s_\ell(x_2 y_0 - x_0 y_2)}{x_2-x_0}. & \end{matrix}$$

The s_ℓ, $\ell = 0, 1$, are *free parameters* whose magnitude will be determined shortly.

The iterates of the operator T applied to f_0 generate a sequence of continuous functions:

$$f_{k+1} = Tf_k = T(T^k f_0), \qquad k \in \mathbb{N}. \tag{2.29}$$

It can be shown that if $\max\{|s_0|, |s_1|\} < 1$, this sequence of continuous functions $\{f_k\}$ converges to a *continuous* function f as $k \to \infty$:

$$f_k(x) \to f(x), \quad \text{for all } x \in [x_0, x_2] \text{ as } k \to \infty.$$

with the property that $f(x_j) = y_j$, $j = 0, 1, 2$. The limit function f is called an affine fractal (interpolation) function [1, 12]. The term *fractal* expresses the, in general, *fractal nature* of the graph of f.

It follows from Eqn. (2.29) that f is the *unique* fixed point of the operator T:

$$\boxed{Tf = f \Longleftrightarrow f(x) = \Lambda(x) + \textstyle\sum_{\ell=0}^{1} s_\ell f\left(\frac{x-b_\ell}{a_\ell}\right).} \tag{2.30}$$

Here

$$\Lambda(x) = \begin{cases} \lambda_0\left(\frac{x-b_0}{a_0}\right), & x_0 \le x \le x_1 \\ \lambda_1\left(\frac{x-b_1}{a_1}\right), & x_1 < x \le x_2. \\ 0, & \text{otherwise} \end{cases}$$

and f was set to be identically zero outside $[0, 1]$.

Eqn. (2.30) also implies that the graph of f is made up of two affine images of itself, each of which is made up of two affine images of itself, each of which is ... ad infinitum! Generally, an affine fractal function does not have a closed representation, i.e., it is not possible to write such a function in terms of simple expressions.

The *recursive* structure of Eqn. (2.30) allows the exact calculation of *moments* and *inner products* of affine fractal functions. To this end, let f be an affine fractal function. The *zeroth moment of f* is defined as

$$M_0(f) = \int_{x_0}^{x_2} f(x)dx, \tag{2.31}$$

and the *first moment of f* by

$$M_1(f) = \int_{x_0}^{x_2} x\, f(x)dx. \tag{2.32}$$

Employing Eqn. (2.30), it can be shown that

$$\boxed{M_0(f) = \frac{\sum_{\ell=0}^{1} a_\ell \int_{x_0}^{x_2} \lambda_\ell(x)dx}{1-\sum_{\ell=0}^{1} a_\ell\, s_\ell}} \tag{2.33}$$

and

$$\boxed{M_1(f) = \frac{\sum_{\ell=0}^{1}\left[a_\ell \int_{x_0}^{x_2}(a_\ell x+b_\ell)\,\lambda_\ell(x)dx+a_\ell\, b_\ell\, s_\ell\, M_0(f)\right]}{1-\sum_{\ell=0}^{1} a_\ell^2\, s_\ell}.} \tag{2.34}$$

These results together with Eqn. (2.30) then provide a formula for the L^2-inner product of two fractal functions f and g interpolating at the same knot points $\{x_0, x_1, x_2\}$.

$$\boxed{\langle f, g\rangle = \frac{\sum_{\ell=0}^{1}\langle\lambda_\ell,\mu_\ell\rangle+\sum_{\ell=0}^{1} a_\ell\, s_\ell[\langle\lambda_\ell,g\rangle+\langle\mu_\ell,f\rangle]}{1-\sum_{\ell=0}^{1} a_\ell\, s_\ell^2},} \tag{2.35}$$

where the μ_ℓ are the affine functions associated with g. The inner product $\langle\lambda_\ell, g\rangle$ can be expressed as

$$\boxed{\langle\lambda_\ell, g\rangle = a_\ell\, M_1(g) + b_\ell\, M_0(g),} \tag{2.36}$$

and similarly for $\langle\mu_\ell, f\rangle$. In other words, to calculate any of the preceding quantities only the *interpolation points* and the free parameter s_0 and s_1 need to be known; they alone completely determine the a_ℓ, b_ℓ, c_ℓ, and d_ℓ.

Example 2.3 *Let $x_0 = 0$, $x_1 = 1/2$, and $x_2 = 1$. Let $s_0 = s_1 = -1/5$. Then*

$$\lambda_0(x) = x \qquad \lambda_1(x) = -x + 1.$$

The fixed point equation *(2.30) of the resulting affine fractal function u then reads*

$$u(x) = \begin{cases} 2x - (1/5)u(2x), & 0 \le x \le 1/2 \\ 2x - 1 - (1/5)u(2x-1), & 1/2 \le x \le 1 \end{cases}$$

Notice that, replacing x by $x/2$ in the above equation, gives precisely Eqn. (2.26) with $s_0 = s_1 = -1/5$.

As pointed out above, the second component of the GHM element, v, consists of two fractal functions; the first interpolates $(-1, 0)$, $(-1/2, -3/10)$, *on the interval* $[-1, 0]$*, and the second* $(0, 1)$*,* $(1/2, -3/20)$*, on the interval* $[0, 1]$*. The affine functions* λ_ℓ *are thus given by*

$$\text{On } [-1,0]: \qquad \lambda_0(x) = -x/10 - 1/10 \quad \lambda_1(x) = 3x/2 + 6/5, \tag{2.37}$$

and

$$\text{On } [0,1]: \qquad \lambda_0(x) = -3x/2 + 6/5 \quad \lambda_1(x) = x/10 - 1/10, \tag{2.38}$$

respectively.

Despite their "jagged nature", affine fractal functions do possess a certain degree of *regularity.* This degree of regularity depends on the size of the parameters s_0 and s_1. For simplicity, assume that $a_0 = a_1 = 1/2$. An *affine* fractal function belongs to the Sobolev space $H^s(\mathbb{R})$ [14] if

$$\boxed{2^{2t-1}\left(s_0^2 + s_1^2\right) < 1.} \tag{2.39}$$

and $s < \min\{t, 3/2\}$.

Example 2.4 *For the affine fractal functions* u *and* v *in the GHM element, condition (2.39) is satisfied for all* $s < 3/2$*. In particular, this implies that* u *and* v *have derivatives everywhere on* $[0, 1]$*, respectively* $[-1, 1]$*, except at points of the form* $i/2^j$*,* $i = 0, 1, \ldots, 2^j$*,* $j = 0, 1, \ldots$*, respectively* $i/2^j$*,* $i = -2^j, \ldots, -1, 0, 1, \ldots, 2^j$*,* $j = 0, 1, \ldots$*.*

The derivative of an affine fractal function f obeys a similar fixed point equation as (2.30):

$$\boxed{f'(x) = \begin{cases} (1/a_0)\lambda_0'\left(\frac{x-b_0}{a_0}\right) + (1/a_0)\, s_0 f'\left(\frac{x-b_0}{a_0}\right), & x_0 < x < x_1 \\ (1/a_1)\lambda_1'\left(\frac{x-b_1}{a_1}\right) + (1/a_1)\, s_1 f'\left(\frac{x-b_1}{a_1}\right), & x_1 < x < x_2 \end{cases}} \tag{2.40}$$

Notice that Eqn. (2.40) is just (2.30) with λ_ℓ replaced by $(1/a_\ell)\lambda_\ell'$ and s_ℓ by s_ℓ/a_ℓ. This observation, in particular, implies that the formulas for the moments and inner products apply to derivatives of affine fractal functions as well (with the above replacements.) These formulas are explicitly given by

$$\boxed{M_0(f') = \frac{\sum_{\ell=0}^{1} \int_{x_0}^{x_2} \lambda_\ell'(x)dx}{1 - \sum_{\ell=0}^{1} s_\ell},} \tag{2.41}$$

and

$$\boxed{\langle f', g' \rangle = \frac{\sum_{\ell=0}^{1}(1/a_\ell)\int_{x_0}^{x_2}\lambda_\ell'(x)\mu_\ell'(x)dx + \sum_{\ell=0}^{1}(s_\ell/a_\ell)\left[\int_{x_0}^{x_2}\lambda_\ell'(x)g'(x)dx + \int_{x_0}^{x_2}\mu_\ell'(x)f'(x)dx\right]}{1-\sum_{\ell=0}^{1}(s_\ell^2/a_\ell)}.} \tag{2.42}$$

2.6 Multiresolution Analyses

Refinable FIS spaces are examples of sequences of *nested spaces* generating multiresolution analyses. More precisely, a *multiresolution analysis (MRA) of $L^2(\mathbb{R})$ (of multiplicity r)*, consists of a sequence of closed subspaces $\{V_k : k \in \mathbb{N}_0\}$, $\mathbb{N}_0 = \{0, 1, 2, \ldots\}$, of $L^2(\mathbb{R})$ [7] such that

Nestedness $V_k \subset V_{k+1}$, for all $k \in \mathbb{N}_0$.

Approximation The union $\bigcup_{k\in\mathbb{N}_0} V_k$ is dense in $L^2(\mathbb{R})$.

Shift-invariance The spaces V_k are shift-invariant, i.e., $f(x) \in V_k$ implies $f(x-\ell) \in V_k$, for all ℓ in some index set I_k.

Refinability $f(x) \in V_k$ if and only if $f(2x) \in V_{k+1}$, for all $k \in \mathbb{N}_0$.

Basis Property The space V_0 is generated by a finite set of functions $\Phi = \{\phi^1, \phi^2, \ldots, \phi^r\}$: $V_0 = \sigma[\Phi]$. In other words, the functions $\phi^1, \phi^2, \ldots, \phi^r$ are the generators of V_0; every $f \in V_0$ may be written in the form

$$f(x) = \sum_{\ell\in I_k} \boldsymbol{c}_\ell^T \Phi(2x-\ell), \quad \Phi(x) = \begin{pmatrix} \phi^1(x) \\ \vdots \\ \phi^r(x) \end{pmatrix}, \tag{2.43}$$

for a *unique* sequence of real (column) vectors $\{\boldsymbol{c}_\ell : \ell \in I_k\}$.

Stability The collection Φ of functions is *uniformly stable*, i.e., there exist positive constants R_1 and R_2 such that for any sequence of real matrices $\{\boldsymbol{C}_\ell : \ell \in I_k\}$

$$\boxed{0 < R_1 \|\boldsymbol{C}_\ell\|_{\ell^2} \leq \|\textstyle\sum_{\ell\in I_k} \boldsymbol{C}_\ell \Phi(\cdot - \ell)\|_{L^2} \leq R_2 \|\boldsymbol{C}_\ell\|_{\ell^2}.} \tag{2.44}$$

Here, $\|\boldsymbol{C}_\ell\|_{\ell^2} = \sum_{\ell\in I_k} \|\boldsymbol{C}_\ell\|^2$.

[7] A subspace S of $L^2(\mathbb{R})$ is *closed* if the limit of every convergent sequence of functions defined on S is in S.

Remarks 2.1 *1. If $\Phi = \{\phi^1, \phi^2, \dots, \phi^r\}$ are uniformly stable generators for V_0, then $\Phi_k = \{\phi^1(2^k x), \phi^2(2^k x), \dots, \phi^r(2^k x)\}$ are uniformly stable generators for V_k. (This follows from refinability).*

2. *The nestedness of the spaces V_k implies the existence of real matrices $\{\boldsymbol{G}_{k\ell}\}$ such that*

$$\boxed{\Phi(2^k x - \ell) = \textstyle\sum_{\ell' \in I_{k+1}} \boldsymbol{G}_{k,\ell'-2\ell} \Phi(2^{k+1} x - \ell'),} \tag{2.45}$$

for all $\ell \in I_k$. Refinability implies that the index k in $\boldsymbol{G}_{k\ell}$ concerns only the size of $\boldsymbol{G}_{k\ell}$, but not the entries. *The matrices $\{\boldsymbol{G}_{k\ell}\}$ are sometimes called the* matrix mask *or the* low pass filter mask *of the scaling vector Φ.*

3. *Eqn. (2.45) may be rewritten in matrix form as follows. Define $\boldsymbol{\Phi}_k(x) = (\Phi(2^k x - \ell) : \ell \in I_k)$ and similarly for $\boldsymbol{\Phi}_{k+1}(x)$. Then*

$$\boxed{\boldsymbol{\Phi}_k = \boldsymbol{A}_k^T \, \boldsymbol{\Phi}_{k+1},} \tag{2.46}$$

where $\boldsymbol{A}_k^T$ is the $\#I_k \times \#I_{k+1}$ matrix whose entries are $\boldsymbol{A}_{k,\ell\ell'} = \boldsymbol{G}_{k,\ell'-2\ell}$.

4. *For computational purposes, one usually requires the generators of V_0 to be* compactly supported. *By refinability, this then implies that the generators for all the spaces V_k are compactly supported.*

An MRA is called *orthogonal* if the generators $\Phi = \{\phi^1, \phi^2, \dots, \phi^r\}$ of V_0, and thus the generators Φ_k of V_k, are orthogonal. In the context of *vector functions*, orthogonality if defined as follows. Two (column) vector functions $\Xi = (\xi_1 \xi_2 \dots \xi_r)^T$ and $\Theta = (\vartheta_1 \vartheta_2 \dots \vartheta_r)^T$ are orthogonal on $L^2(\Omega)$ if

$$\int_\Omega \Xi(x)\Theta^T(x)dx = \begin{pmatrix} \int_\Omega \xi_1(x)\vartheta_1(x)dx & \cdots & \int_\Omega \xi_1(x)\vartheta_r(x)dx \\ \int_\Omega \xi_2(x)\vartheta_1(x)dx & \cdots & \int_\Omega \xi_2(x)\vartheta_r(x)dx \\ \vdots & \cdots & \vdots \\ \int_\Omega \xi_r(x)\vartheta_1(x)dx & \cdots & \int_\Omega \xi_r(x)\vartheta_r(x)dx \end{pmatrix} = \boldsymbol{O}_{r\times r}. \tag{2.47}$$

Here $\boldsymbol{O}_{r\times r}$ denotes the $r \times r$ zero matrix. To simplify notation, the inner product of vector functions Ξ and Θ on $L^2(\Omega)$ will also be written as $\langle \Xi, \Theta \rangle$.

Remark 2.2 *The (column) vectors in Eqn. (2.43) are the inner products of f with the vector function Φ:*

$$\boldsymbol{c}_\ell^T = \langle f, \Phi(2x-\ell)\rangle = \begin{pmatrix} \int_\Omega f(x)\phi^1(2x-\ell)dx \\ \vdots \\ \int_\Omega f(x)\phi^r(2x-\ell)dx \end{pmatrix}^T \tag{2.48}$$

The generators $\Phi = (\phi^1, \ldots, \phi^r)^T$ of an MRA are usually termed *multiscaling functions* or *scaling vector.* If $r = 1$, the generator is called a *scaling function.*

An example of an orthogonal MRA of $L^2(\mathbb{R})$ of multiplicity 2 is provided by the GHM element [9] constructed above. The generators are clearly compactly supported and continuous. In the case $r = 1$, the *Daubechies scaling functions* [8] are a family of compactly supported orthogonal generators with increasing regularity for an MRA of $L^2(\mathbb{R})$. The simplest scaling function in this family is the *Haar scaling function* ϕ^H. It is defined by

$$\phi^H(x) = \begin{cases} 1, & 0 \le x \le 1 \\ 0, & \text{otherwise.} \end{cases} \tag{2.49}$$

Since ϕ^H is supported on the interval $[0,1]$ it is automatically orthogonal to its integer translates. Moreover,

$$\phi^H(x) = \frac{1}{2}\phi^H(2x) + \frac{1}{2}\phi^H(2x-1), \tag{2.50}$$

i.e., $g_0 = 1/2 = g_1$. The spaces V_k which are generated by the Haar scaling function $\{\phi^H(2^k x)\}$ consist of all piecewise constant functions from $L^2(\mathbb{R})$ supported on intervals of the form $[\ell/2^k, (\ell+1)/2^k]$, $\ell \in \mathbb{Z}$.

As $V_k \subset V_{k+1}$, the *difference spaces* $V_{k+1} \ominus V_k$, consisting of all those functions in V_{k+1} which are *not* in V_k may be employed to obtain a *multiscale basis* for $L^2(\mathbb{R})$. These difference spaces are usually called *(multi)wavelet spaces* and denoted by W_k:

$$\boxed{V_{k+1} = V_k \oplus W_k, \qquad \text{for all } k \in \mathbb{N}_0.} \tag{2.51}$$

The generators of W_k are called *(multi)wavelets.* The Ψ_k must be chosen in such a way that the $\{\phi_k^1, \phi_k^2, \ldots, \phi_k^r\} \cup \{\psi_k^1, \psi_k^2, \ldots, \psi_k^r\}$ are uniformly stable. If the spaces V_k have r generators, the spaces W_k have also r generators

$\Psi_k = \{\psi_k^1, \psi_k^2, \ldots, \psi_k^r\}$. Furthermore, $\psi_k^i(x) = \psi^i(2^k x)$, $i = 1, \ldots, r$, where ψ^i is one of the r generators of W_0. If the generators of V_0, and thus of V_k, are orthogonal the sum $V_k \oplus W_k$ is an orthogonal sum; every $f \in V_k$ is orthogonal to every $g \in W_k$: $\langle f, g\rangle = 0$. Since $W_k \subset V_{k+1}$, there exist real matrices $\{\boldsymbol{H}_{k\ell} : \ell \in J_k\}$ such that

$$\boxed{\Psi(2^k x - \ell) = \textstyle\sum_{\ell' \in J_k} \boldsymbol{H}_{k,\ell'-2\ell}\Phi(2^{k+1}x - \ell'),} \tag{2.52}$$

As before, the index k on $\boldsymbol{H}_{k\ell}$ indicates only the size of the matrix, the entries are independent of k. Here J_k is an index set such that $\#I_{k+1} = \#I_k + \#J_k$. Commonly used terms for the matrices $\{\boldsymbol{H}_{k\ell}\}$ are *matrix mask* or *high pass filter mask.* As above, it is convenient to define $\boldsymbol{\Psi}_k(x) = (\Psi(2^k x - \ell) : \ell \in J_k)$. Then

$$\boxed{\boldsymbol{\Psi}_k = \boldsymbol{B}_k^T \boldsymbol{\Phi}_{k+1},} \tag{2.53}$$

where $\boldsymbol{B}_k^T$ is the $\#J_k \times \#J_{k+1}$ matrix whose entries are $\boldsymbol{B}_{k,\ell\ell'} = \boldsymbol{H}_{k,\ell'-2\ell}$.

One important feature of scaling vectors and multiwavelets is that all relevant information about their properties is contained in the matrices $\{\boldsymbol{G}_\ell\}$ and $\{\boldsymbol{H}_\ell\}$, respectively. In particular, one has

$$\boxed{\textstyle\sum_{\ell \in \mathbb{Z}} \boldsymbol{G}_\ell \boldsymbol{G}_{\ell-2m}^T = 2\delta_m \boldsymbol{I}_{r\times r},} \tag{2.54}$$

(Orthogonality of integer translates for Φ)

$$\boxed{\textstyle\sum_{\ell \in \mathbb{Z}} \boldsymbol{H}_\ell \boldsymbol{H}_{\ell-2m}^T = 2\delta_m \boldsymbol{I}_{r\times r},} \tag{2.55}$$

(Orthogonality of integer translates for Ψ)

$$\boxed{\textstyle\sum_{\ell \in \mathbb{Z}} \boldsymbol{G}_\ell \boldsymbol{H}_{\ell-2m}^T = \boldsymbol{O}_{r\times r},} \tag{2.56}$$

(Orthogonality between Φ and Ψ)

$$\boxed{\textstyle\sum_{\ell \in \mathbb{Z}} \boldsymbol{G}_{m-2\ell}^T \boldsymbol{G}_{n-2\ell} + \boldsymbol{H}_{m-2\ell}^T \boldsymbol{H}_{n-2\ell} = \delta_{mn}\boldsymbol{I}_{r\times r}, \quad m, n \in \mathbb{Z},} \tag{2.57}$$

($V_0 + W_0 = V_1$)

Here δ_m and δ_{mn} are the *Kronecker Deltas*:

$$\delta_m = \begin{cases} 1, & m = 0 \\ 0, & \text{otherwise} \end{cases}, \qquad \delta_{mn} = \begin{cases} 1, & m = n \\ 0, & \text{otherwise} \end{cases}.$$

Given an orthogonal MRA generated by a scaling vector Φ, one way of obtaining orthogonal generators for the multiwavelet spaces is by choosing the matrices $\{\boldsymbol{H}_\ell : \ell \in J_0\}$ so that Eqns. (2.55) – (2.57) are satisfied. In the case when $r = 1$, the *scalars* h_ℓ have a particularly simple dependence on the g_ℓ:

$$\boxed{h_\ell = (-1)^\ell g_{N-\ell}.} \tag{2.58}$$

Here N is one less than the number of terms in the refinement equation for the Daubechies scaling function ϕ: $\phi(x) = \sum_{\ell=0}^{N} g_\ell \phi(2x - \ell)$.

For example, the wavelet associated with the Haar scaling function has $h_0 = 1/2$ and $h_1 = -1/2$ and is explicitly expressible as

$$\psi^H(x) = \begin{cases} 1 & 0 \le x \le 1/2 \\ -1 & 1/2 \le x \le 1 \\ 0 & \text{otherwise.} \end{cases} \tag{2.59}$$

The multiwavelet associated with the GHM element consists of two generators both of which are affine fractal (interpolation) functions. Figure 2 below displays their graphs. Note that it is possible to have one generator symmetric and the other one antisymmetric about the y-axis. This is an important feature of the GHM and DGHM element. In the Daubechies family there *do not* exist symmetric/antisymmetric scaling functions or wavelets, with the exception of the discontinuous Haar scaling function and wavelet. Certain applications in image compression and signal processing do require symmetric/antisymmetric bases.

As already indicated above, the GHM and DGHM element does not have closed representation. This also holds true for the Daubechies family of scaling functions and wavelets; the only exception is the discontinuous Haar scaling function and wavelet. However, for all computational purposes, it is *not* necessary to have an explicit expression for scaling functions, scaling vectors, wavelets, or multiwavelets. All the information about these functions is contained in the *known and accessible* matrices $\{\boldsymbol{G}_{k\ell}\}$ and $\{\boldsymbol{H}_{k\ell}\}$. In

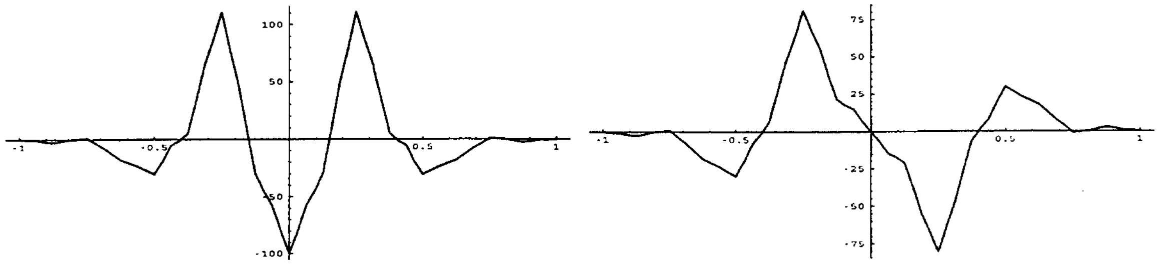

Figure 2: The orthogonal DGHM multiwavelet.

order to elaborate on this point, the fundamental concept of a *multiscale transformation* needs to be introduced.

Convention: Unless explicitly mentioned, all scaling vectors and the multiwavelets are assumed to be orthogonal generators of their respective spaces.

Since V_{k+1} is a direct sum of V_k and W_k, the $\#I_{k+1} \times \#I_{k+1}$ matrix $\boldsymbol{M}_k = \begin{pmatrix} \boldsymbol{A}_k^T \\ \boldsymbol{B}_k^T \end{pmatrix}$ is invertible and

$$\begin{pmatrix} \boldsymbol{\Phi}_k \\ \boldsymbol{\Psi}_k \end{pmatrix} = \boldsymbol{M}_k \, \boldsymbol{\Phi}_{k+1}. \tag{2.60}$$

Denote the inverse of $\boldsymbol{M}_k$ by $\boldsymbol{W}_k = (\, \boldsymbol{C}_k^T \quad \boldsymbol{D}_k^T \,)$. Then

$$\boxed{\boldsymbol{\Phi}_{k+1} = \boldsymbol{C}_k^T \, \boldsymbol{\Phi}_k + \boldsymbol{D}_k^T \, \boldsymbol{\Psi}_k.} \tag{2.61}$$

The matrices $\boldsymbol{C}_k^T$ and $\boldsymbol{D}_k^T$ satisfy

$$\boxed{\boldsymbol{A}_k \, \boldsymbol{C}_k + \boldsymbol{B}_k \, \boldsymbol{D}_k = \boldsymbol{I}_{k+1},} \tag{2.62}$$

and

$$\boxed{\boldsymbol{C}_k \boldsymbol{A}_k = \boldsymbol{I}_k, \quad \boldsymbol{D}_k \boldsymbol{B}_k = \boldsymbol{I}_k^*, \quad \boldsymbol{D}_k \boldsymbol{A}_k = \boldsymbol{O} = \boldsymbol{C}_k \boldsymbol{B}_k.} \tag{2.63}$$

(Compare with Eqns. (2.54) – (2.57)!) Here $\boldsymbol{I}_{k+1}$ denotes the $\#I_{k+1} \times \#I_{k+1}$ identity matrix, $\boldsymbol{I}_k$ the $\#I_k \times \#I_k$ identity matrix, and $\boldsymbol{I}_k^*$ the $\#J_k \times \#J_k$ identity matrix. Equations (2.62) and (2.63) are referred to as *multifilter relations.* If the generators for V_k and W_k are *orthonormal* then $\boldsymbol{M}_k^{-1} = \boldsymbol{M}_k^T$. The matrix $\boldsymbol{W}_k$ is also called the *discrete multiwavelet transform.*

Example 2.5 *Consider the Haar scaling function ϕ^H and the Haar wavelet ψ^H. The functions $\boldsymbol{\Phi}_k$ and $\boldsymbol{\Psi}_k$ are given by*

$$\boldsymbol{\Phi}_k(x) = (2^{k/2}\phi^H(2^k x - \ell) : \ell = 0, 1, \ldots, 2^k - 1),$$

and

$$\boldsymbol{\Psi}_k(x) = (2^{k/2}\psi^H(2^k x - \ell) : \ell = 0, 1, \ldots, 2^k - 1),$$

respectively. Thus, $I_k = \{0, 1, \ldots, 2^k - 1\} = J_k$. The factor $2^{k/2}$ was added to normalize *the functions: $\int_0^1 |2^{k/2}\phi^H(2^k x - \ell)|^2 dx = \int_0^1 |2^{k/2}\psi^H(2^k x - \ell)|^2 dx = 1$. Then, since*

$$\phi^H(x) = \frac{1}{2}\phi^H(2x) + \frac{1}{2}\phi^H(2x-1),$$
$$\psi^H(x) = \frac{1}{2}\phi^H(2x) - \frac{1}{2}\phi^H(2x-1),$$

one has

$$\phi^H(2^k x - \ell) = 2^{-1/2}[\phi^H(2^{k+1}x - 2\ell) + \phi^H(2^{k+1}x - 2\ell - 2)]$$

and

$$\psi^H(2^k x - \ell) = 2^{-1/2}[\phi^H(2^{k+1}x - 2\ell) - \phi^H(2^{k+1}x - 2\ell - 2)].$$

Thus, the $2^{k+1} \times 2^k$ *matrices* $\boldsymbol{A}_k$ *and* $\boldsymbol{B}_k$ *are given by*

$$\boldsymbol{A}_k = \begin{pmatrix} \frac{1}{\sqrt{2}} & 0 & 0 & 0 & \cdots & 0 \\ \frac{1}{\sqrt{2}} & 0 & 0 & 0 & \cdots & 0 \\ 0 & \frac{1}{\sqrt{2}} & 0 & 0 & \cdots & 0 \\ 0 & \frac{1}{\sqrt{2}} & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & 0 & 0 & \frac{1}{\sqrt{2}} & 0 \\ 0 & \cdots & 0 & 0 & \frac{1}{\sqrt{2}} & 0 \\ 0 & \cdots & 0 & 0 & 0 & \frac{1}{\sqrt{2}} \\ 0 & \cdots & 0 & 0 & 0 & \frac{1}{\sqrt{2}} \end{pmatrix}$$

and

$$\boldsymbol{B}_k = \begin{pmatrix} \frac{1}{\sqrt{2}} & 0 & 0 & 0 & \cdots & 0 \\ -\frac{1}{\sqrt{2}} & 0 & 0 & 0 & \cdots & 0 \\ 0 & \frac{1}{\sqrt{2}} & 0 & 0 & \cdots & 0 \\ 0 & -\frac{1}{\sqrt{2}} & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & 0 & 0 & \frac{1}{\sqrt{2}} & 0 \\ 0 & \cdots & 0 & 0 & -\frac{1}{\sqrt{2}} & 0 \\ 0 & \cdots & 0 & 0 & 0 & \frac{1}{\sqrt{2}} \\ 0 & \cdots & 0 & 0 & 0 & -\frac{1}{\sqrt{2}} \end{pmatrix}$$

respectively. Eqn. (2.61) reads for the Haar scaling function and wavelet as follows

$$\phi^H(2^{k+1}x - 2\ell) = 2^{-1/2}[\phi^H(2^k x - \ell) + \psi^H(2^k x - \ell)]$$

and

$$\phi^H(2^{k+1}x - 2\ell - 1) = 2^{-1/2}[\phi^H(2^k x - \ell) - \psi^H(2^k x - \ell)].$$

2.7 Multiscale bases and multiscale transformations

Starting with a fixed level $k = K$ in the decomposition (2.51) of V_{k+1} into V_k and W_k and proceeding to level 0, one obtains a *multiscale representation* of a function $f \in V_K$. More precisely,

$$\boxed{V_K = V_0 \bigoplus_{k=0}^{K-1} W_k,} \tag{2.64}$$

and a function $f \in V_K$ may be expressed in *single-scale representation* with respect to the *single-scale basis* $\{\Phi_K\}$ as

$$f(x) = \sum_{\ell \in I_K} c_{K,\ell} \Phi_K(x - \ell) = \boldsymbol{c}_K^T \boldsymbol{\Phi}_K, \tag{2.65}$$

or in *multiscale representation* with respect to the *multiscale basis* $\{\Phi_0\} \bigcup_{k=0}^{K-1} \{\Psi_k\}$ as

$$\begin{aligned} f(x) &= \sum_{\ell \in I_0} c_{0,\ell} \Phi_0(x-\ell) + \sum_{\ell \in J_0} d_{0,\ell} \Psi_0(x-\ell) + \cdots + \sum_{\ell \in J_{K-1}} d_{K-1,\ell} \Psi_{K-1}(x-\ell) \\ &= \boldsymbol{c}_0^T \boldsymbol{\Phi}_0 + \sum_{k=0}^{K-1} \boldsymbol{d}_k^T \boldsymbol{\Psi}_k. \end{aligned} \tag{2.66}$$

Here the c and d are real scalar coefficients whereas $\boldsymbol{c}$ and $\boldsymbol{d}$ are real vector coefficients whose dimension is given by the cardinality of the corresponding index set. Recall that $\Phi_k(x - \ell) = \Phi(2^k x - \ell)$ and likewise for $\Psi_k(x - \ell)$. Employing Eqn. (2.61), the multiscale representation of $\boldsymbol{\Phi}_k$ may be succinctly written as

$$\boxed{\boldsymbol{\Phi}_K = \left(\prod_{k=0}^{K-1} \boldsymbol{W}_{K-k}\right) \begin{pmatrix} \boldsymbol{\Phi}_0 \\ \boldsymbol{\Psi}_0 \\ \vdots \\ \boldsymbol{\Psi}_{K-1} \end{pmatrix},} \tag{2.67}$$

where the $I_k \times I_k$ matrix $\boldsymbol{W}_k$ is of the form

$$\boldsymbol{W}_k = \begin{pmatrix} \boldsymbol{W}_{k-1} & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{I} \end{pmatrix}, \tag{2.68}$$

with $\boldsymbol{O}$ being the zero matrix and $\boldsymbol{I}$ the $J_{k-1} \times J_{k-1}$ identity matrix, respectively.

The matrix $\left(\prod_{k=0}^{K-1} \boldsymbol{W}_{K-k}\right)$ is an example of a *multiscale transformation.* A multiscale transformation applied to the vector coefficients of a function $f \in V_K$ gives rise to the so-called *fast reconstruction* and decomposition algorithm. To this end, consider

$$\begin{aligned} \boldsymbol{c}_{k+1}^T \boldsymbol{\Phi}_{k+1} &= \boldsymbol{c}_k^T \boldsymbol{\Phi}_k + \boldsymbol{d}_k^T \boldsymbol{\Psi}_k, && (2.69) \\ &= (\boldsymbol{c}_k^T \boldsymbol{A}_k^T + \boldsymbol{d}_k^T \boldsymbol{B}_k^T) \boldsymbol{\Phi}_{k+1}. && (2.70) \end{aligned}$$

Hence,

$$\boxed{\boldsymbol{c}_{k+1} = \boldsymbol{A}_k\,\boldsymbol{c}_k + \boldsymbol{B}_k\,\boldsymbol{d}_k = \boldsymbol{M}_k^T \begin{pmatrix} \boldsymbol{c}_k \\ \boldsymbol{d}_k \end{pmatrix}.} \tag{2.71}$$

In other words, $\boldsymbol{M}_k^T$ is an operator which maps the pair of sequences $(\boldsymbol{c}_k, \boldsymbol{d}_k)$ of length $\#I_k$ and $\#J_k$, respectively, to the single sequence $\boldsymbol{c}_{k+1}$ of length $\#I_{k+1}$.

Remark 2.3 *In the present setting, namely an MRA on $L^2(\mathbb{R})$, $\#J_k = \#I_k = (1/2)\#I_{k+1}$. This situation changes when the real line $\mathbb{R}$ is replaced by a proper open subset Ω.*

In component notation, Eqn. (2.71) reads

$$\boxed{\boldsymbol{c}_{k+1,\ell} = \textstyle\sum_{\ell'} \boldsymbol{G}_{\ell'-2\ell}\,\boldsymbol{c}_{k,\ell} + \boldsymbol{H}_{\ell'-2\ell}\,\boldsymbol{d}_{k,\ell}.} \tag{2.72}$$

Note that only the *even* indices are used to obtain $\boldsymbol{c}_k$. Zeros are used for the odd indices (interlacing of zeros or upsampling by 2: $\uparrow 2$).

The reconstruction algorithm (2.71) is schematically described in Figure 3. The matrices $\boldsymbol{A}_k$ act along the horizontal arrows, whereas the matrices

$$\boxed{\begin{array}{ccccccccc} \boldsymbol{c}_0 & \longrightarrow & \boldsymbol{c}_1 & \longrightarrow & \cdots & \longrightarrow & \boldsymbol{c}_{K-1} & \longrightarrow & \boldsymbol{c}_K \\ & \nearrow & & \nearrow & & \nearrow & & \nearrow & \\ \boldsymbol{d}_0 & & \boldsymbol{d}_1 & \cdots & \boldsymbol{d}_{K-2} & & \boldsymbol{d}_{K-1} & & \end{array}} \tag{2.73}$$

Figure 3: The reconstruction algorithm

$\boldsymbol{B}_k$ act along the diagonal arrows.

Reversing the reconstruction algorithm produces the *decomposition algorithm.* Multipliy Eqn. (2.69) by $\boldsymbol{\Phi}_k^T$ and use the orthogonality relations between $\boldsymbol{\Phi}_k$ and $\boldsymbol{\Psi}_k$ to get

$$\boldsymbol{c}_{k+1}^T \langle \boldsymbol{\Phi}_{k+1}, \boldsymbol{\Phi}_k \rangle = \boldsymbol{c}_k^T \langle \boldsymbol{\Phi}_k, \boldsymbol{\Phi}_k \rangle;$$

now employ Eqn. (2.46) and transpose to obtain

$$\boxed{\boldsymbol{c}_k = \boldsymbol{A}_k^T\,\boldsymbol{c}_{k+1}.} \tag{2.74}$$

In a similar way one shows that

$$\boxed{\boldsymbol{d}_k = \boldsymbol{B}_k^T \, \boldsymbol{c}_{k+1}.} \tag{2.75}$$

In component notation these two equations read

$$\boxed{\boldsymbol{c}_{k,\ell} = \sum_{\ell'} \boldsymbol{G}_{\ell'-2\ell} \, \boldsymbol{c}_{k+1\ell} \quad \text{and} \quad \boldsymbol{d}_{k,\ell} = \sum_{\ell'} \boldsymbol{H}_{\ell'-2\ell} \, \boldsymbol{c}_{k\ell}.} \tag{2.76}$$

Notice that $\boldsymbol{c}_k$ and $\boldsymbol{d}_k$ are *sampled* only at the *even* integers (down-sampling by 2: $\downarrow 2$).

Combining Eqns. (2.74) and (2.75) into one gives

$$\boxed{\begin{pmatrix} \boldsymbol{c}_k \\ \boldsymbol{d}_k \end{pmatrix} = \begin{pmatrix} \boldsymbol{A}_k^T \\ \boldsymbol{B}_k^T \end{pmatrix} \boldsymbol{c}_{k+1} = \boldsymbol{M}_k \, \boldsymbol{c}_{k+1}.} \tag{2.77}$$

Hence, the operator M_k assigns to each input sequence $\boldsymbol{c}_{k+1}$ of length $\#I_{k+1}$ a pair of output sequences of length $\#I_k$ and $\#J_k$, respectively, where $\#I_{k+1} = \#I_k + \#J_k$. The matrices $\boldsymbol{A}_k^T$ act along the horizontal arrows, the $\boldsymbol{B}_k^T$ along

$$\boxed{\begin{array}{ccccccccc} \boldsymbol{c}_0 & \longrightarrow & \boldsymbol{c}_1 & \longrightarrow & \cdots & \longrightarrow & \boldsymbol{c}_{K-1} & \longrightarrow & \boldsymbol{c}_K \\ & \searrow & & \searrow & & \searrow & & \searrow & \\ & & \boldsymbol{d}_0 & & \boldsymbol{d}_1 & \cdots & \boldsymbol{d}_{K-2} & & \boldsymbol{d}_{K-1} \end{array}} \tag{2.78}$$

Figure 4: The decomposition algorithm

the downward diagonals.

If the scaling vector Φ and the associated multiwavelet Ψ have short support, then the matrices $\boldsymbol{M}_k$ and $\boldsymbol{W}_k$ are *uniformly sparse*. If, in addition, $\#I_{k+1}/\#I_k \geq \varrho > 1$ (in the present setting, $\varrho = 2$; cf. Remark (2.3)), then application of $\prod_{k=0}^{K-1} \boldsymbol{W}_{K-k}$ and $(\prod_{k=0}^{K-1} \boldsymbol{W}_{K-k})^{-1}$ requires the order of $\#I_K$ operations, uniformly in K [3].

Letting $K \to \infty$ in Eqn. (2.66) gives the *resolution of* $L^2(\Omega)$ in terms of the *sequence of multiwavelet spaces* $\{W_k : k \in \mathbb{N}_0\}$:

$$\boxed{L^2(\Omega) = V_0 \oplus \bigoplus_{k=0}^{\infty} W_k.} \tag{2.79}$$

Hence, any f in $L^2(\Omega)$ may be represented in the form

$$\boxed{f(x) = \boldsymbol{c}_0^T \boldsymbol{\Phi}_0 + \sum_{k=0}^{\infty} \boldsymbol{d}_k^T \boldsymbol{\Psi}_k(x) = \langle f, \boldsymbol{\Phi}_0 \rangle \, \boldsymbol{\Phi}_0 + \sum_{k=0}^{\infty} \langle f, \boldsymbol{\Psi}_k \rangle \, \boldsymbol{\Psi}_k(x).} \tag{2.80}$$

The $\boldsymbol{d}_k$ are called the *multiwavelet coefficients of f.*

3 Approximation-theoretic Issues

The focus of this section is on the approximation and regularity properties of Galerkin-type methods. In such methods, the solution space, usually some Hilbert space $\mathcal{H}$, is approximated by an ascending sequence of *trial* or *approximation spaces* $\{V_k : k = 0, 1, \ldots\}$. The quality of the approximation as well as the regularity of the solution depend on approximation- theoretic properties of the trial spaces.

3.1 Jackson and Bernstein estimates

The two basic estimates that give the *quality of approximation*, respectively, the *regularity* of the *approximant* are

- *Direct* or *Jackson Estimate*:

$$\boxed{\inf\{\|f - v\|_{L^2} : v \in V_k\} \leq C_1 \, 2^{-sk} \|f\|_{H^s}, \qquad f \in H^s(\Omega).} \tag{3.1}$$

  The positive constant C_1 is *independent* of k and s, but may depend on Ω.

- *Indirect* or *Bernstein Estimate*:

$$\boxed{\|v\|_{H^s} \leq C_2 \, 2^{sk} \, \|v\|_{L^2}, \qquad v \in V_k.} \tag{3.2}$$

  Again, the positive constant C_2 is *independent* of k and s, but may depend on Ω.

Remarks 3.1 *1. Eqn. (3.1) estimates the error in approximating a function f in the Sobolev space $H^s(\Omega)$ by an element v of the trial space V_k. Note that the smoother f or the larger k, the better the approximation. The exponent $-sk$ gives the* rate of approximation *or* rate of convergence *of the* approximator v *to the approximant f as $k \to \infty$.*

2. *Eqn. (3.2) estimates the regularity of the approximator v in terms of its L^2-norm.*

3. *Closely related to the Jackson estimate is the concept of* approximation order*[4, 11, 10]. Let $h > 0$ be an integer, and let $v_h(x) = v(x/h)$. Define V_h to be the space consisting of all functions v_h such that $v \in V$ (V_h is a dilate by h of V). A space V has* approximation order s *if for every compactly supported function $f \in H^s(\mathbb{R})$*

$$\boxed{\inf\{\|f - v\|_{L^2} : v \in V\} \leq C_3\, h^s \|f\|_{H^s},} \tag{3.3}$$

for some positive constant C_3 independent of h and s. It is well-known that a space V has approximation order s if and only if it contains the space of real polynomials $\mathbb{P}^s$ of order at most s, i.e., of degree at most $s-1$: $\mathbb{P}^s \subseteq V$. Note that s here is necessarily an integer ≥ 0.

It is known that the refinable space $V = \sigma[h]$ generated by the hat function h has approximation order 2. In other words, every function in $V = \sigma[h]$ reproduces exactly polynomials of order at most two, i.e., of degree at most one: $\mathbb{P}^2 \subseteq V = \sigma[h]$. Recall that the last statement means that there exists a finite sequence of real constants c_ℓ with the property that

$$x^p = \sum_\ell c_\ell h(x - \ell), \qquad p = 0, 1$$

on any *finite* interval with integer endpoints.

Example 3.1 *It was shown in [6] that the GHM element has the same approximation order as the hat function, namely two. (Cf. Eqn. (2.27)) Using the interpolatory nature of the GHM element, one can easily derive the vector constants in Eqn. (2.27). For instance, the* partition of unity *for the DGHM is obtained as follows. Let $\chi(x) = \begin{cases} 1, & 0 \leq x \leq 1 \\ 0, & otherwise \end{cases}$ and let $\phi^l(x) = \phi^2(x)$[8] restricted to the interval $[0,1]$: $\phi^l(x) = \begin{cases} \phi^2(x), & 0 \leq x \leq 1 \\ 0, & otherwise \end{cases}$, and let $\phi^r(x) = \phi^2(x-1)$ restricted to $[0,1]$: $\phi^r(x) = \begin{cases} \phi^2(x-1), & 0 \leq x \leq 1 \\ 0, & otherwise \end{cases}$. Set*

$$\chi(x) = 1 = c_1\phi^l(x) + c_2\phi^1(x) + c_3\phi^r(x), \quad \text{for all } 0 \leq x \leq 1.$$

[8] To adhere to standard notation, we set $\phi^1 = u$ and $\phi^2 = v$.

Now use that $\phi^l(0) = 1$, $\phi^1(0) = 0 = \phi^r(0)$, $\phi^l(1/2) = \phi^r(1/2) = -3/10$, $\phi^r(1) = 1$, *and* $\phi^l(1) = 0 = \phi^1(0)$, *to obtain*

$$c_1 = 1; \qquad c_2 = 8/5 \qquad c_3 = 1,$$

and thus

$$\begin{aligned} \chi(x) &= \phi^l(x) + (8/5)\phi^1(x) + \phi^r(x) \\ &= \begin{pmatrix} 8/5 & 1 \end{pmatrix} \begin{pmatrix} \phi^1(x) \\ \phi^2(x) \end{pmatrix} + \begin{pmatrix} 0 & 1 \end{pmatrix} \begin{pmatrix} \phi^1(x-1) \\ \phi^2(x-1) \end{pmatrix}. \end{aligned}$$

To obtain the vector coefficients for the reproduction of, say the function $\rho(x) = \begin{cases} x, & 0 \le x \le 1 \\ 0, & otherwise \end{cases}$, *the orthogonality properties of the GHM element may be employed. For this purpose, let*

$$\rho(x) = 1 = c_1\phi^l(x) + c_2\phi^1(x) + c_3\phi^r(x), \quad \textit{for all } 0 \le x \le 1.$$

and multiply this equation by ϕ^l, ϕ^r, *and* ϕ^1, *respectively, and integrate over the interval* $[0, 1]$. *By orthogonality, this yields*

$$\begin{aligned} \int_0^1 \rho(x)\phi^l(x)dx &= c_1 \int_0^1 \phi^l(x)\phi^l(x)dx, \\ \int_0^1 \rho(x)\phi^r(x)dx &= c_2 \int_0^1 \phi^r(x)\phi^r(x)dx, \\ \int_0^1 \rho(x)\phi^1(x)dx &= c_3 \int_0^1 \phi^1(x)\phi^1(x)dx, \end{aligned}$$

or more succinctly,

$$c_1 = \frac{M_0(\phi^l)}{\|\phi^l\|_{L^2}} \quad c_2 = \frac{M_0(\phi^r)}{\|\phi^r\|_{L^2}} \quad c_3 = \frac{M_0(\phi^1)}{\|\phi^1\|_{L^2}}.$$

Employing Eqns. (2.34) and (2.35), gives

$$c_1 = 0, \qquad c_2 = 3/5, \qquad c_3 = 1.$$

That c_1 *needs to be zero should be clear since* $\phi^l(0) = 1$, $\phi^r(0) = 0 = \phi^1(0)$, *and* $\rho(0) = 0$.

The next result indicates under what conditions a scaling vector generating an MRA satisfies the fundamental Jackson- and Bernstein-type estimates. To this end, let $\mathbf{\Phi}_k$ be a collection of functions from $L^2(\Omega)$. **If**

Orthogonality: $\mathbf{\Phi}_k$ is an *orthogonal* collection of functions, i.e., $\langle \mathbf{\Phi}_k, \mathbf{\Phi}_k \rangle = \boldsymbol{O}$.

Uniform Boundedness: The functions $\phi_{k,\ell}$ in $\mathbf{\Phi}_k$ satisfy $\|\phi_{k,\ell}\|_{L^2} \leq C$, for all k and ℓ and some positive constant C *independent of k and ℓ.*

Compact Support: The functions in $\mathbf{\Phi}_k$ are compactly supported.

Approximation: The space V generated by $\mathbf{\Phi}_k$ contains the real polynomials of order s: $\mathbb{P}^s \subseteq V = \sigma[\mathbf{\Phi}_k]$.

Regularity: Let γ be the largest positive number so that for all $s < \gamma$, each function $\phi_{k,\ell}$ in $\mathbf{\Phi}_k$ is in $H^s(\Omega)$.

then the Jackson estimate for f in $H^s(\Omega)$

$$\boxed{\|f - \langle f, \mathbf{\Phi}_k \rangle \mathbf{\Phi}_k\|_{L^2} \leq c_1 \, 2^{-sk} \, \|f\|_{H^s}.} \tag{3.4}$$

and the Bernstein estimate

$$\boxed{\|v\|_{H^s} \leq c_2 \, 2^{sk} \, \|v\|_{L^2}, \qquad v \in V = \sigma[\mathbf{\Phi}_k].} \tag{3.5}$$

hold. The inner product in Eqn. (3.4) are the vector coefficients in the expansion of a function f with respect to the orthogonal generators $\mathbf{\Phi}_k$ of $V = \sigma[\mathbf{\Phi}_k]$, and, as usual, c_1 and c_2 are positive constants independent of k and s.

Suppose the sequence $\{V_k : k \in \mathbb{N}_0\}$ generates an *orthogonal* MRA of $L^2(\Omega)$ and $\mathbf{\Phi}_k$ is a basis for V_k. For a function $f \in L^2(\Omega)$, let P_k denote the *projection of f onto V_k*. Then $P_k f$ may be written in the form

$$\boxed{P_k f(x) = \boldsymbol{c}_k^T \mathbf{\Phi}_k = \textstyle\sum_{\ell \in I_k} \boldsymbol{c}_{k\ell}^T \Phi(2^k x - \ell),} \tag{3.6}$$

for vector coefficients $\boldsymbol{c}_k$ and $\boldsymbol{c}_{k\ell}$. (In case that V_0 is generated by only *one* scaling function, the coefficients $\boldsymbol{c}_{k\ell}$ are scalars.) Using the orthogonality of the scaling vectors $\mathbf{\Phi}_k$, Eqn. (3.6) may be written as

$$\boxed{P_k f(x) = \langle f, \mathbf{\Phi}_k \rangle \, \mathbf{\Phi}_k.} \tag{3.7}$$

(Cf. Remark (2.2)).

Remark 3.1 *The orthogonality of the scaling vector $\mathbf{\Phi}$ is equivalent to*

$$\boxed{P_k P_m = P_k, \quad \textit{for all } k \leq m.} \tag{3.8}$$

Since the multiwavelet spaces are given by $W_k = V_{k+1} \ominus V_k$, the difference $(P_{k+1} - P_k)f$ is an element of W_k, and by orthogonality, each function in W_k may be written in this form. The orthogonal projections $Q_{k+1} = P_{k+1} - P_k$ project a function f from $L^2(\mathbb{R})$ onto the multiwavelet space W_k. Moreover,

$$P_k f = \sum_{j=0}^{k} (P_j - P_{j-1}) f = \sum_{j=0}^{k} Q_j f. \tag{3.9}$$

Here, $P_{-1} = 0$.

For an orthogonal MRA, one can show that if the Jackson estimate (3.4) and the Bernstein estimate (3.5) holds, then

$$\boxed{C_1 \, ||f||^2_{H^s} \leq \textstyle\sum_{k=0}^{\infty} 2^{2sk} ||(P_k - P_{k-1})f||^2_{L^2} \leq C_2 \, ||f||^2_{H^s},} \tag{3.10}$$

for positive constants C_1 and C_2. (P_{-1} is defined as 0.) Note that $2^{2sk}||(P_k - P_{k-1})f||^2_{L^2}$ is a *weighted* L^2-inner product on the multiwavelet spaces W_k. The *regularity* of the projection $P_k f$ of f is bounded by the regularity of f:

$$\boxed{||P_k f||_{H^s} \leq C_3 ||f||_{H^s},} \tag{3.11}$$

for some positive constant C_3. Notice that by the orthogonality of the multiwavelet spaces, $\sum_{k=0}^{\infty} 2^{2sk} ||(P_k - P_{k-1})f||^2_{L^2} = \left|\sum_{k=0}^{\infty} 2^{sk}(P_k - P_{k-1})f\right|^2_{L^2}$. In terms of the multiwavelet bases, the inequalities in (3.10) read

$$\boxed{C_1 \, ||f||_{H^s} \leq ||\boldsymbol{D}_k^s \langle f, \mathbf{\Psi}\rangle^T ||_{\ell^2} \leq C_2 \, ||f||_{H^s}.} \tag{3.12}$$

Here, $\boldsymbol{D}_k^s$ denotes a *diagonal matrix* with entries 2^{ks} and $\mathbf{\Psi}$ the *entire* multiwavelet basis including $\mathbf{\Phi}_0$. The inequalities in (3.12) express the fact that the *weighted wavelet coefficients* are indicative of the regularity of the function f.

3.2 Preconditioning

Here it is assumed that the spaces $\{V_k : k \in \mathbb{N}_0\}$ are used for trial spaces in a Galerkin method. Then the solution u of the operator equation $Lu = f$ is approximated by $u_k = P_k u = \sum_{j=0}^{k}(P_j - P_{j-1})u = \sum_{j=0}^{k} Q_j u$, i.e., its (orthogonal) projection onto V_k. If a basis is chosen in such a way that the conditions stated in the previous subsection are fulfilled, then the Jackson estimate (with $f = u$ and $u_k = \langle u, \boldsymbol{\Phi}_k \rangle \boldsymbol{\Phi}_k$) gives the *rate of convergence* of the approximate solution u_k to the exact solution u: $u_k \to u$ as fast as $2^{-ks} \to 0$ as $k \to \infty$. In other words, this rate is determined by the *degree of regularity* and the *approximation order* of the solution u. The efficiency of a Galerkin methods depends on

- The regularity/approximation order of the trial spaces V_k;
- The condition number and sparsity of the stiffness matrix.

Example 3.2 *The GHM element has approximation order 2 and smoothness $s < \gamma = 3/2$. Thus, the rate of convergence is governed by $s < \gamma = 3/2$. This rate is* the same *at that for linear finite elements, i.e., the (non-orthogonal) spaces generated by the hat function h.*

To simplify notation, let $\Sigma^s f = \sum_{k=0}^{\infty} 2^{sk}(P_k - P_{k-1})f = \sum_{k=0}^{\infty} 2^{sk} Q_k f$. Notice that this may be thought of as a *weighted representation* of f in the multiwavelet spaces W_k. In terms of a basis, $Q_k f$ is given by

$$Q_k f = \langle f, \boldsymbol{\Psi}_k \rangle \, \boldsymbol{\Psi}_k. \tag{3.13}$$

Using this notation, the inequalities in (3.12) show that Σ^s, if interpreted as an operator on the function f, is a *shift* in the *scale of Sobolev spaces*: If f is in H^t then $\Sigma^s f$ is in H^{s+t}. (The statements in (3.12) only apply to $t = 0$, i.e., $H^0 = L^2$, but the result holds for all $t \neq 0$.) Now the operator L also acts as a shift in the scale of Sobolev spaces: If L has order t and u is in H^s then Lu is in H^{s-t}. The idea behind preconditioning is to undo the effect of L onto u by applying a shift Σ^t. For this purpose, consider the stiffness matrix $\boldsymbol{S}_k$ *expressed in the multiwavelet basis* $\{\boldsymbol{\Phi}_0\} \bigcup_{j=0}^{k-1}\{\boldsymbol{\Psi}_j\}$ *of* V_k. To simplify notation, let $\boldsymbol{\Phi}_0 = \boldsymbol{\Psi}_{-1}$ and write $\boldsymbol{\Psi}$ for the entire multiwavelet basis $\bigcup_{j=-1}^{k-1}\{\boldsymbol{\Psi}_j\}$. The stiffness matrix is explicitly given by

$$\boldsymbol{S}_k = \langle L\boldsymbol{\Psi}, \boldsymbol{\Psi} \rangle^T. \tag{3.14}$$

It is known that $\boldsymbol{S}_k$ is ill-conditioned for operators L of degree $t \neq 0$; *the condition number of* $\boldsymbol{S}_k$ *increases exponentially in* k (cf., for instance, [3, 7, 15, 16, 17]). The following line of reasoning shows how a preconditioner is obtained. To this end, a simplification in notation is needed. We write $\|f\|_{\mathcal{H}_1} \sim \|f\|_{\mathcal{H}_2}$ instead of $c_1\|f\|_{\mathcal{H}_1} \leq \|f\|_{\mathcal{H}_2} \leq c_2\|f\|_{\mathcal{H}_1}$.

For the sake of simplicity, assume that L is a *linear self-adjoint elliptic operator of order* $2s$. Under this assumption, one has (cf. also (2.8))

$$\boxed{\|P_k^* L u_k\|_{H^{-s}} \sim \|u_k\|_{H^s}, \qquad u_k \in V_k.} \tag{3.15}$$

(Here, * denotes the *adjoint*). For the approximate solution $u_k \in V_k$, let $w_k = \Sigma^s u_k$. Then by (3.10), for $s \neq 0$,

$$\|w_k\|_{L^2} = \|\Sigma^s u_k\|_{L^2} \sim \|u_k\|_{H^s},$$

and by (3.15)

$$\|u_k\|_{H^s} \sim \|P_k^* L u_k\|_{H^{-s}}.$$

Again apply (3.10), but now to $\|P_k^* L u_k\|_{H^{-s}}$, to obtain

$$\|P_k^* L u_k\|_{H^{-s}} \sim \|(\Sigma^{-s})^* P_k^* L P_k \Sigma^{-s} w_k\|_{L^2},$$

and hence,

$$\|w_k\|_{L^2} \sim \|(\Sigma^{-s})^* P_k^* L P_k \Sigma^{-s} w_k\|_{L^2}. \tag{3.16}$$

This last equivalence means that the operator $L_k^s = (\Sigma^{-s})^* P_k^* L P_k \Sigma^{-s}$ have degree zero (it maps L^2 into L^2) and are thus *uniformly boundedly invertible*, that is,

$$\|L_k^s\|_{L^2},\ \|(L_k^s)^{-1}\|_{L^2} = \mathcal{O}(1) \quad \text{as } k \to \infty.$$

For, the right hand inequality in

$$0 < c_1\|w_k\|_{L^2} \leq \|L_k^s w_k\|_{L^2} \leq c_2\|w_k\|_{L^2}$$

gives uniform boundedness of L_k^s, whereas the right hand inequality yields uniform boundedness for $(L_k^s)^{-1}$.

In particular, the matrix representation of L_k^s with respect to the multiwavelet basis $\boldsymbol{\Psi}$, given by

$$\boxed{\langle L_k^s \boldsymbol{\Psi}, \boldsymbol{\Psi}\rangle^T = \boldsymbol{D}_k^{-s}\, \boldsymbol{S}_k\, \boldsymbol{D}_k^{-s}} \tag{3.17}$$

has *uniformly bounded condition number:* cond $(\boldsymbol{D}_k^{-s})\,\boldsymbol{S}_k\,\boldsymbol{D}_k^{-s}) = \mathcal{O}(1)$. Recall that $\boldsymbol{D}_k^{-s}$ is the diagonal matrix whose entries are 2^{-sk}.

The (linear) algebraic system associated with the operator equation $Lu = f$ has preconditioner $\boldsymbol{P}_k = \boldsymbol{D}_k^{-s}$ and is of the form

$$\boldsymbol{P}_k\,\boldsymbol{S}_k\,\boldsymbol{P}_k \boldsymbol{d}_k = \boldsymbol{P}_k\,\boldsymbol{f}_k.$$

In the above equation, $\boldsymbol{d}_k$ are the multiwavelet coefficients of u_k with respect to the basis $\boldsymbol{\Psi}_k$ and $\boldsymbol{f}_k = \langle f, \boldsymbol{\Psi}_k \rangle$.

Remarks 3.2 *1. The stiffness matrix $\boldsymbol{S}_k$ with respect to the entire multiwavelet basis $\boldsymbol{\Psi}$ may be obtained from the stiffness matrix $\langle L\boldsymbol{\Phi}_k, \boldsymbol{\Phi}_k \rangle$ with respect to the* single scale*:*

$$\boxed{\boldsymbol{S}_k = \boldsymbol{W}_k^T \langle L\boldsymbol{\Phi}_k, \boldsymbol{\Phi}_k \rangle \boldsymbol{W}_k.} \tag{3.18}$$

2. The preconditioner $\boldsymbol{P}_k$ is essentially a change of basis; it relates the single scale representation to the multiscale representation as reflected by the weighted operator Σ^s. Based on this interpretation an efficient algorithm for the computation of $\boldsymbol{P}_k\,\boldsymbol{S}_k\,\boldsymbol{P}_k$ employing Eqn. (3.18) may be obtained (cf. [3]).

(a) Compute $\boldsymbol{y} = \boldsymbol{W}_k \boldsymbol{D}_k^{-s} \boldsymbol{x}$. The structure of the multiwavelet transform $\boldsymbol{W}_k$ and the geometrical growth of the number $\#\boldsymbol{\Phi}_k$ of the basis elements $\boldsymbol{\Phi}_k$ as k increases allows this computation to be of order $\#\boldsymbol{\Phi}_k$.

(b) Compute $\boldsymbol{z} = \langle L\boldsymbol{\Phi}_k, \boldsymbol{\Phi}_k \rangle \boldsymbol{y}$. The sparseness of $\langle L\boldsymbol{\Phi}_k, \boldsymbol{\Phi}_k \rangle$ requires operations of order $\#\boldsymbol{\Phi}_k$.

(c) Compute $\boldsymbol{D}_k^{-s} \boldsymbol{W}_k^T \boldsymbol{z}$. This is essentially (a) above.

Therefore, the number of operations required and the memory allocated for the process is of order $\#\boldsymbol{\Phi}_k$.

4 A Simple Example: $-\Delta u = f$

In this section, the simple differential equation $-\Delta u = f$ on $\Omega = [0, 1]$ with boundary conditions $u(0) = 0 = u(1)$ is considered. A Galerkin method based on a *single scale representation* is employed to obtain an approximate solution to the differential equation. The presentation of this example is of a more didactical than mathematical nature; nevertheless, this example contains all the ingredients and addresses all the relevant issues needed for a more sophisticated *multiscale* approach to linear second order elliptic differential equations.

Following Example (2.1), the natural setting for a Galerkin method is to employ as trial and test spaces subspaces of the Sobolev space $H_0^1([0,1])$. For this purpose, appropriate local bases for the ascending sequence of such trial and test spaces need to be constructed. In this section, the GHM element $\Phi = (\phi^1\ \phi^2)^T$, as defined in Section 2, and the orthogonal MRA generated by it is used to obtain such trial and test spaces. Since the test and trial functions are to vanish at the boundary of $[0,1]$, we define define

$$\boxed{V_0^\circ = \sigma[\phi^1],} \tag{4.1}$$

and for $k \geq 1$,

$$\boxed{V_k^\circ = \sigma[\mathbf{\Phi}_k],} \tag{4.2}$$

where $\mathbf{\Phi}_k$ is the vector function

$$\boxed{\mathbf{\Phi}_k(x) = \begin{pmatrix} \phi^1(2^k x) \\ \phi^1(2x-1) \\ \vdots \\ \phi^1(2^k x - [2^k - 1]) \\ \phi^2(2^k x - 1) \\ \vdots \\ \phi^2(2^k x - [2^k - 1]) \end{pmatrix}} \tag{4.3}$$

Hence, at level $k \geq 1$, scale 2^{-k}, the space V_k° is spanned by the 2^k translates of the function $\phi^1(2^k x)$ and by the 2^{k-1} translates of the function $\phi^2(2^k x - 1)$. Hence,

$$\boxed{\dim V_k^\circ = \#\mathbf{\Phi}_k = 2^{k+1} - 1, \qquad k \geq 1.} \tag{4.4}$$

Remark 4.1 *The nested sequence of spaces* $\{V_k^\circ : k \in \mathbb{N}_0\}$ *generates an* orthogonal *MRA of* $L_0^2([0,1]) = \{f \in L^2([0,1]) : f(0) = 0 = f(1)\}$.

To justify the above remark, some important geometric properties of the GHM element need to be mentioned explicitly.

- By Example (2.4) the functions ϕ^1 and ϕ^2, and thus their dilates and translates are elements of $H^s(\Omega)$ for $s < \gamma = 3/2$.

- By construction, the restriction of $\phi^i(2^k x - \ell)$, $i = 1, 2$, to $\Omega = [0, 1]$ *preserves orthogonality between the translates.* This is a feature not enjoyed by, for instance, the Daubechies scaling functions (except the Haar scaling function). There is *no* need to introduce so-called boundary functions to retain orthogonality between the translates.
  This observation allows the construction of an orthogonal MRA of $L^2([0, 1])$ [6, 12].

- Discarding in the approximation spaces for the orthogonal MRA of $L^2([0, 1])$ all those functions that *do not* vanish at the boundary yields an *orthogonal* MRA of $L_0^2([0, 1])$.

The GHM element fulfills all the conditions listed in the previous section to guarantee the validity of a Jackson and Bernstein estimate. The Jackson estimate, in particular, implies that the rate of convergence of an approximate solution u_k to the exact solution u is given by $s < \gamma = 3/2$.

Now let u_k be the orthogonal projection of the exact solution u onto the trial space V_k° $(k \geq 1)$, i.e.,

$$u_k(x) = \boldsymbol{c}_k^T\, \boldsymbol{\Phi}_k(x), \tag{4.5}$$

for a coefficient vector of length $\#\boldsymbol{\Phi}_k = 2^{k+1} - 1$.

Substituting this expression into $-\Delta u = f$, multiplying by both sides by $\boldsymbol{\Phi}_k^T$, integrating by parts and using the boundary conditions yields

$$\boxed{\langle \boldsymbol{\Phi}_k', \boldsymbol{\Phi}_k' \rangle^T\, \boldsymbol{c}_k = \langle f, \boldsymbol{\Phi}_k \rangle^T.} \tag{4.6}$$

To compute the entries in the stiffness matrix $\boldsymbol{S}_k = \langle \boldsymbol{\Phi}_k', \boldsymbol{\Phi}_m' \rangle^T$, notice that there are three types of innerproducts in $\boldsymbol{S}_k$:

1. Inner products of the form $\int_0^1 (\phi^1(2^x - \ell))' \cdot (\phi^1(2^x - m))'\, dx$.

2. Inner products of the form $\int_0^1 (\phi^2(2^x - \ell))' \cdot (\phi^2(2^x - m))'\, dx$.

3. Inner products of the form $\int_0^1 (\phi^1(2^x - \ell))' \cdot (\phi^2(2^x - m))'\, dx$.

Each of these three types is evaluated by a first applying a change in variables $\bar{x} = 2^k x - \ell$. Taking into account that the support of $\phi^1(2^k x - \ell)$ and $\phi^2(2^k x - \ell)$ is an interval of length 2^{-k}, respectively, $2 \cdot 2^{-k}$ yields

1. $\int_0^1 (\phi^1(2^x - \ell))' \cdot (\phi^1(2^x - m))'\, dx = 2^k \int_0^1 (\phi^1)'(x)(\phi^1)'(x - m + \ell) dx$

$$= \begin{cases} 2^k, & \ell = m \\ 0, & \text{otherwise} \end{cases} .$$

2. $\int_0^1 (\phi^2(2^x - \ell))' \cdot (\phi^2(2^x - m))'\, dx = 2^k \int_0^1 (\phi^2)'(x)(\phi^2)'(x - m + \ell) dx$

$$= \begin{cases} 2^k, & m = \ell, \ell \pm 1 \\ 0, & \text{otherwise} \end{cases} .$$

3. $\int_0^1 (\phi^1(2^x - \ell))' \cdot (\phi^2(2^x - m))'\, dx = 2^k \int_0^1 (\phi^1)'(x)(\phi^2)'(x - m + \ell) dx$

$$= \begin{cases} 2^k, & m = \ell, \ell + 1 \\ 0, & \text{otherwise} \end{cases} .$$

Hence the $(2^{k+1} - 1) \times (2^{k+1} - 1)$ stiffness matrix has the following block form:

$$\langle \mathbf{\Phi}'_k, \mathbf{\Phi}'_k \rangle^T = 2^k \cdot \begin{pmatrix} \boldsymbol{B}_1 & \boldsymbol{B}_2 \\ \boldsymbol{B}_2^T & \boldsymbol{B}_3 \end{pmatrix}, \tag{4.7}$$

where $\boldsymbol{B}_1$ is the $2^k \times 2^k$ *diagonal* matrix

$$\boxed{\boldsymbol{B}_1 = \operatorname{diag}\left(\langle (\phi^1)', (\phi^1)' \rangle\right),} \tag{4.8}$$

$\boldsymbol{B}_2$ the $(2^k - 1) \times 2^k$ matrix

$$\boxed{\boldsymbol{B}_2 = \begin{pmatrix} a & b & 0 & \cdots & 0 \\ 0 & a & b & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & a & b \\ 0 & 0 & \cdots & 0 & a \end{pmatrix}} \tag{4.9}$$

with $a = \langle(\phi^1)', (\phi^2)'(\cdot+1)\rangle$ and $b = \langle(\phi^1)', (\phi^2)'\rangle$, and $\boldsymbol{B}_3$ the *tridiagonal* $(2^k-1)\times(2^k-1)$ matrix

$$\boxed{\boldsymbol{B}_3 = \begin{pmatrix} c & d & 0 & \cdots & 0 \\ d & c & d & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & c & d \\ 0 & 0 & \cdots & d & c \end{pmatrix}} \tag{4.10}$$

with $c = \langle(\phi^2)', (\phi^2)'\rangle$ and $d = \langle(\phi^2)', (\phi^2)'(\cdot-1)\rangle$. Note that the entries in the matrices $\boldsymbol{B}_1$, $\boldsymbol{B}_2$, and $\boldsymbol{B}_3$ are *independent of the level k.*

The entries in the matrices $\boldsymbol{B}_1$, $\boldsymbol{B}_2$, and $\boldsymbol{B}_3$ can be computed *exactly* using Eqns. (2.41) and (2.42). This computation yields

$$\left\langle(\phi^1)', (\phi^1)'\right\rangle = 100/21, \qquad \langle(\phi^2)', (\phi^2)'\rangle = 85/21, \tag{4.11}$$

$$\left\langle(\phi^1)', (\phi^2)'(\cdot+1)\right\rangle = -80/21, \quad \langle(\phi^2)', (\phi^2)'(\cdot-1)\rangle = 43/21. \tag{4.12}$$

The linear system (4.6) may now be solved by applying the preconditioner $\boldsymbol{P}_k$ derived in the previous section.

Remark 4.2 *It must be stressed that the above Galerkin method based on the GHM element has* same *rate of convergence of u_k to the exact solution u as a Galerkin method based on the (nonorthogonal) MRA generated by the hat function.*

References

[1] M. F. Barnsley, "Fractal functions and interpolation," *Constr. Approx.* Vol. 5 No. 1 (1986), 303 – 329.

[2] J. H. Bramble, J. E. Pasciak, and J. Xu, "Parallel multilevel preconditioners," *Math. Comput.* **55** (1990), 1 – 22.

[3] W. Dahmen, "Wavelet and multiscale methods for operator equations," *to appear in Acta Numerica.*

[4] C. de Boor, R. A. DeVore, and A. Ron, "The structure of finitely generated shift-invariant spaces in $L^2(\mathbb{R}^d)$," *J. Funct. Anal.* **119** (1994), 37 – 78.

[5] G. Donavan, J. S. Geronimo, and D. P. Hardin, "Intertwining multiresolution analyses and the construction of piecewise polynomial wavelets," *SIAM J. Math. Anal.* **27** (1996) 1791 – 1815.

[6] G. Donovan, J. S. Geronimo, D. P. Hardin, and P. R. Massopust, "Construction of orthogonal wavelets using fractal interpolation functions," *SIAM J. Math. Anal.* Vol. 27, No. 4 (1996), 1158 – 1192.

[7] W. Dahmen and A. Konuth, "Multilevel preconditioning," *Numer. Math.* **63** (1992), 315 – 344.

[8] I. Daubechies, *Ten Lectures on Wavelets*, SIAM, Vol. 61, Philadelphia, 1992.

[9] J. S. Geronimo, D. P. Hardin, and P. R. Massopust, "Fractal functions and wavelet expansions based on several scaling functions," *J. Approx. Th.* Vol. 78, No. 3 (1994), 373 – 401.

[10] C. Heil, G. Strang, and V. Strela, "Approximation by translates of refinable functions," *Numer. Math.* **73** (1996), 75 – 94.

[11] R.- Q. Jia, "Refinable shift-invariant spaces: From Splines to Wavelets," *Approximation Theory VIII, Vol. 2: Wavelets and Multilevel Approximation* (1995), 179 – 208.

[12] P. R. Massopust, *Fractal Functions, Fractal Surfaces, and Wavelets*, Academic Press, San Diego, 1994.

[13] P. R. Massopust, "A multiwavelet based on piecewise C^1 fractal functions and related applications to differential equations," *Preprint.*

[14] P. R. Massopust, "Fractal functions and applications," *Chaos, Fractals, and Solitons* Vol. 8 No. 2 (1997), 171 – 190.

[15] P. Oswald, "On discrete norm estimates related to multilevel preconditoners in the finite element method," in *Constructive Theory of Functions*, K. G. Ivanov, P. Petrushev, and B. Sendov (eds.), Proc. Int. Conf. Varna 1991, Bulg. Acad. Sci., Sofia (1992), 203 – 214.

[16] P. Oswald, *Multilevel Finite Element Approximation – Theory and Applications*, Teubner Skripten zur Numerik, Teubner, Stuttgart, 1994.

[17] H. Yserentant, "Two preconditioners based on the multilevel splitting of finite element methods," *Numer. Math.* **58** (1990), 163 – 184.

Distribution:

1 MS 0188 C. E. Meyers, 4523
1 MS 0321 W. J. Camp, 9200
1 MS 1111 S. S. Dosanjh, 9221
1 MS 1110 R. C. Allen, Jr., 9205
1 MS 1110 D. E. Womble, 9222
1 MS 1110 D. S. Greenberg, 9223
1 MS 1111 G. S. Heffelfinger, 9225
1 MS 0441 R. W. Leland, 9226
1 MS 0819 M. A. Christon, 9231
1 MS 0835 T. E. Voth, 9113
1 MS 0819 T. G. Trucano, 9231
1 MS 0819 J. R. Weatherby, 9231
1 MS 0825 R. S. Baty, 9115
1 MS 0836 S. P. Burns, 9116
1 MS 0437 S. W. Attaway, 9118
5 P. R. Massopust, Dept. of Mathematics, SHSU, Huntsville, TX 77341
1 MS 1110 D. W. Roach, 9222
1 MS 9018 Central Technical Files, 8940-2
5 MS 0899 Technical Library, 4916
2 MS 0619 Review and Approval Desk For DOE/OSTI, 12690